

UNIVERSIDADE FEDERAL DO PARANA

ROBSON SIPRAKI

O USO DA ESTATÍSTICA PARA FINS DE AVALIAÇÃO EM POLÍTICAS  
EDUCACIONAIS: UM ESTUDO A PARTIR DA PESQUISA TALIS (OCDE, 2009).

CURITIBA

2015

ROBSON SIPRAKI

O USO DA ESTATÍSTICA PARA FINS DE AVALIAÇÃO EM POLÍTICAS  
EDUCACIONAIS: UM ESTUDO A PARTIR DA PESQUISA TALIS (OCDE, 2009)

Dissertação ao Curso de Mestrado do Programa de Pós-Graduação em Educação, Setor de Educação, da Universidade Federal do Paraná, como requisito parcial à obtenção do título de Mestre em Educação.

Orientadora: Profa. Dra. RoseMeri Trojan

Co-orientadora: Profa. Dra. Suely Ruiz Giolo

CURITIBA

2015

---

Catálogo na publicação  
Mariluci Zanela – CRB 9/1233  
Biblioteca de Ciências Humanas e Educação - UFPR

Sipraki, Robson

O uso da estatística para fins de avaliação em políticas educacionais: um estudo a partir da pesquisa TALIS (OCDE, 2009) / Robson Sipraki – Curitiba, 2015.

133 f.

Orientadora: Profa. Dra. Rose Meri Trojan

Co-orientadora: Profa. Dra. Suely Ruiz Giolo

Dissertação (Mestrado em Educação) – Setor de Ciências Humanas da Universidade Federal do Paraná.

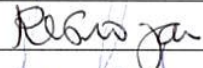
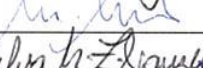
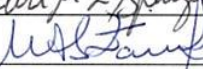

1. Educação - Brasil. 2. Avaliação educacional - Brasil. 3. TALIS. 4. Estatística - Educação. I. Título.

CDD 371.26

## PARECER

Defesa de Dissertação de **Robson Sipraki** para obtenção do Título de MESTRE EM EDUCAÇÃO. Os abaixo assinados, Prof.<sup>a</sup> Dr.<sup>a</sup> Rose Meri Trojan, Prof.<sup>a</sup> Dr.<sup>a</sup> Maria Angélica Pedra Minhoto, Prof.<sup>a</sup> Dr.<sup>a</sup> Nelva Maria Zibetti Sganzerla, Prof.<sup>a</sup> Dr.<sup>a</sup> Maria Amélia Sabbag Zainko, arguiram, nesta data, o candidato acima citado, o qual apresentou a seguinte Dissertação: "O USO DA ESTATÍSTICA PARA FINS DE AVALIAÇÃO EM POLÍTICAS EDUCACIONAIS: UM ESTUDO A PARTIR DA PESQUISA TALIS (OCDE, 2009)".

Procedida a arguição, segundo o Protocolo aprovado pelo Colegiado, a Banca é de Parecer que o candidato está Apto ao Título de MESTRE EM EDUCAÇÃO, tendo merecido as apreciações abaixo:

| BANCA   | ASSINATURA   | APRECIAÇÃO |
|---|--|------------|
| Prof. <sup>a</sup> Dr. <sup>a</sup> Rose Meri Trojan              |  | aprovado   |
| Prof. <sup>a</sup> Dr. <sup>a</sup> Maria Angélica Pedra Minhoto  |  | aprovado   |
| Prof. <sup>a</sup> Dr. <sup>a</sup> Nelva Maria Zibetti Sganzerla |  | Aprovado   |
| Prof. <sup>a</sup> Dr. <sup>a</sup> Maria Amélia Sabbag Zainko    |  | aprovado   |

Curitiba, 27 de março de 2015



Prof.<sup>a</sup>. Dr.<sup>a</sup>. Monica Ribeiro da Silva  
Coordenadora do PPGE

Prof.<sup>a</sup>. Dra. Monica Ribeiro da Silva  
Coordenadora do Programa de  
Pós-Graduação em Educação  
Matrícula: 125750



Dedico esse trabalho aos meus avós Francisca e Silvio e ao meu pai Oliveiros, que partiram para o mundo espiritual, mas que continuam me apoiando. A minha querida e amada mãe Ivone que me amparou em todos os momentos de alegria e tristeza, minha amada irmã Silvana e meu cunhado Francisco que me apoiaram e me fortaleceram nessa caminhada, aos meus sobrinhos queridos João e Beatriz, e minha luz e razão de minha vida meu filho Pedro que enche de alegria e me dá forças nos meus dias.

## **AGRADECIMENTOS**

Eu tive a sorte de encontrar um verdadeiro anjo, que me inspirou, lapidou, ensinou, foi além de professora. O verdadeiro orientador é aquele que mostra os erros e os acertos e explica o porquê. Você, Rose Meri Trojan, é o anjo que ilumina minha vida, esse trabalho não é só meu, é NOSSO! Agradeço a Deus todos os dias pelo privilégio de ter conhecido uma pessoa como você.

O mesmo digo à minha co-orientadora Suely Ruiz Giolo, que me acompanha desde a graduação, outro anjo iluminado que Deus colocou no meu caminho, que me apoiou e acreditou que sou capaz.

Aos funcionários do Colégio Conselheiro Zacarias e especialmente ao diretor Profº Paulo Soares, pelo apoio e compreensão em todos os momentos.

À Vanisse Simone Correia e Gizele Pereira, pelo apoio e o incentivo que dispensaram para comigo do início ao término desse trabalho.

Aos grandes anjos da educação, Professoras Maria Angélica Minhoto, Maria Amélia Sabbag Zainko, e Nelva Maria Zibetti Sganzerla, pelas contribuições e sugestões no trabalho. E também a Regina Maria Micheloto, que me ajudou, incentivou, contribuiu, esteve do meu lado a todo o momento e no final fez uma esplêndida revisão.

Obrigado Deus por ter colocado verdadeiros anjos em minha vida.

Dizer de felicidade vai  
de encontro com tudo  
que amamos,  
que sonhamos,  
que lutamos,  
que vivemos,  
agora e  
para sempre.

*(Robson Sipraki)*

## RESUMO

Esse trabalho tem como objeto o uso da estatística como ferramenta utilizada nas pesquisas educacionais, tomando como foco central o emprego de medidas de tendência central e de dispersão na Pesquisa Internacional sobre Ensino e Aprendizagem (TALIS), realizada pela Organização de Cooperação e Desenvolvimento Econômico (OCDE), por meio de um estudo dos resultados apresentados pelo Brasil, no relatório publicado em 2009. O estudo busca demonstrar o modo de aplicação das medidas de tendência central e de dispersão na pesquisa, a fim de apontar a necessidade de averiguar a adequação ou inadequação das estatísticas utilizadas em pesquisas de larga escala, em âmbito mundial e nacional e dos impactos que poderão acarretar nas tomadas de decisão relacionadas às políticas educacionais. A definição do tema nasceu das necessidades geradas pela pesquisa *O aporte das pesquisas internacionais de avaliação do ensino e da aprendizagem na educação brasileira e as políticas públicas de formação e trabalho docente*, desenvolvida pelo Grupo de Estudos e Pesquisas em Educação Comparada de Políticas Educacionais (GEPEC). Esta investigação justifica-se pelo fato de a estatística se encontrar cada vez mais presente no âmbito da pesquisa em políticas educacionais. Estudar estas questões implica em entender o emprego da estatística no campo da educação, com a finalidade de medir a qualidade dos sistemas educativos e seus desempenhos. O objetivo é avaliar o uso de medidas de tendência central e de dispersão na pesquisa TALIS, com vistas a subsidiar a avaliação e implementação de políticas educacionais relacionadas às condições de trabalho docente. Os procedimentos metodológicos realizados nessa pesquisa tomam como ponto de partida um estudo teórico sobre as finalidades de pesquisas sobre avaliação de políticas educacionais que utilizam dados numéricos para compreender e avaliar a realidade da educação, bem como os fundamentos, conceitos e métodos da Estatística, que se coloca como uma das principais ferramentas.

Palavras-chave: políticas educacionais; avaliação de larga escala; estatística; TALIS.



## **ABSTRACT**

This work has as object the use of statistics as a tool used in educational research, taking as its central focus the use of central tendency and dispersion. To assess the adequacy these employed measures, taking as an example the International Research on Teaching and Learning (TALIS), conducted by the Organization for Economic Cooperation and Development (OECD), through a study of the results presented by Brazil in the report published in 2009. The relevance of this research is justified because the statistical finding increasingly present within the research in educational policies. Studying these issues involves understanding the use of statistics in the field of education, used in large-scale research on a global and national levels and the impacts that may result in decision making related to educational policies. The theme definition arose from the needs generated by the research The contribution of international research evaluation of teaching and learning in Brazilian education and public policy formation and teaching work, developed by the Group of Studies and Research in Education Policy Comparative (GEPEC ). Therefore, the objective is to evaluate the use of statistics in research, in order to support the evaluation and implementation of educational policies with reference to the TALIS survey. The methodological procedures performed in this research take as their starting point a theoretical study about the research purposes of evaluation of educational policies that use numeric data to understand and assess the reality of education as well as the fundamentals, concepts and methods of Statistics, which poses as a major tool.

Keywords: educational policies; large-scale assessment; Statistics; TALIS.

## LISTA DE QUADROS

|  |    |
|--|----|
| Quadro 1: VANTAGENS E DESVANTAGENS DAS RESPOSTAS ABERTAS E FECHADAS.....   | 48 |
| Quadro 2: CARACTERÍSTICAS PARA DIFERENCIAR POPULAÇÃO E AMOSTRA .....   | 59 |
| Quadro 3: PESQUISA HIPOTÉTICA PARA EXEMPLIFICAR O COEFICIENTE DE VARIAÇÃO .....  | 76 |
| Quadro 4: PERGUNTA DO QUESTIONÁRIO DOS PROFESSORES DA PESQUISA TALIS (2007-08) RELACIONADO A IDADE.....                                    | 87 |
| Quadro 5: PERGUNTA DO QUESTIONÁRIO DOS PROFESSORES DAS PESQUISA TALIS (2007-08) RELACIONADO AO TEMPO DE TRABALHO DOCENTE.....              | 88 |
| Quadro 6: PERGUNTA DO QUESTIONÁRIO DOS PROFESSORES DA PESQUISA TALIS (2007-08) REFERENTE ÀS PRÁTICAS DE ENSINO, CRENÇAS E ATITUDES.....    | 91 |
| Quadro 7: PERGUNTA DO QUESTIONÁRIO DOS PROFESSORES DA PESQUISA TALIS (2007-08), REFERENTE À TURMA ESPECÍFICA. ....                         | 95 |
| Quadro 8: PERGUNTA DO QUESTIONÁRIO DOS PROFESSORES DA PESQUISA TALIS (2007-08), REFERENTE À TURMA ESPECÍFICA AOS RECUSRSOS DA ESCOLA. .... | 97 |

## LISTA DE FIGURAS

|                                       |    |
|---------------------------------------|----|
| Figura 1: COMPOSIÇÃO DO SAEB.....     | 30 |
| Figura 2: PIRÂMIDE DA DEFINIÇÃO ..... | 62 |

## LISTA DE TABELAS

|   |     |
|---|-----|
| Tabela 1: IDADE DE UM GRUPO DE PESSOAS .....  | 72  |
| Tabela 2: PERCENTUAL DE PROFESSORES RELACIONADOS À FAIXA ETÁRIA<br>.....  | 88  |
| Tabela 3: PERCENTUAL DE PROFESSORES RELACIONADOS AO REGIME E<br>EXPERIÊNCIA DE TRABALHO (2007-08) .....   | 89  |
| Tabela 4: AUTOEFICÁCIA: PROGRESSO COM OS ALUNOS (2007-08) .....   | 92  |
| Tabela 5: QUESTÃO 31: AUTOEFICÁCIA: SUCESSO COM OS ALUNOS (2007-08)<br>.....  | 93  |
| Tabela 6: QUESTÃO 31: AUTOEFICÁCIA: FAZENDO A DIFERENÇA<br>SIGNIFICATIVA EDUCACIONAL DOS ALUNOS (2007-08).....                                  | 94  |
| Tabela 7: QUESTÃO 43: CLIMA DISCIPLINAR EM SALA DE AULA: BARULHO EM<br>SALA DE AULA (2007-08) .....   | 95  |
| Tabela 8: QUESTÃO 30: EM QUE MEDIDA O ENSINO ESTÁ SENDO<br>PREJUDICADO .....  | 96  |
| Tabela 9: COMPARAÇÃO ENTRE VARÁVEIS DA QUESTÃO Nº 30 -<br>QUESTIONÁRIO DO PROFESSOR.....  | 99  |
| Tabela 10: NÚMERO DE ALUNOS MATRICULADOS POR ESCOLA DO BANCO DE<br>DADOS OCDE 2009. ....  | 100 |
| Tabela 11: MÉDIA DE ALUNOS POR TURMA DO RELATÓRIO DA OCDE 2009.   | 101 |
| Tabela 12: DESCRIÇÃO DO NÚMERO DE ALUNOS POR TURMA DO BANCO DE<br>DADOS OCDE. ....  | 102 |
| Tabela 13: DESCRIÇÃO DOS DADOS RELACIONADOS À COMPARAÇÃO<br>MÉDIAS DE ALUNOS POR TURMA ENTRE BRASIL E ESPANHA<br>PARA O TESTE PARAMÉTRICO ..... | 105 |
| Tabela 14: ANÁLISE DA COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA<br>ENTRE BRASIL E ESPANHA, PARA O TESTE PARAMÉTRICO.....                         | 106 |

|  |     |
|--|-----|
| Tabela 15: DESCRIÇÃO DOS DADOS RELACIONADOS À COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE BRASIL E MÉXICO, PARA O TESTE PARAMÉTRICO.....          | 107 |
| Tabela 16: ANÁLISE DA COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE BRASIL E MÉXICO PARA O TESTE PARAMÉTRICO.....                                   | 107 |
| Tabela 17: DESCRIÇÃO DOS DADOS RELACIONADOS À COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE ESPANHA E ITÁLIA, PARA O TESTE PARAMÉTRICO.....         | 108 |
| Tabela 18: ANÁLISE DE COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE ESPANHA E ITÁLIA, PELO TESTE PARAMÉTRICO.....                                   | 109 |
| Tabela 19: DESCRIÇÃO DOS DADOS RELACIONADOS À COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE BRASIL E ESPANHA PARA O TESTE NÃO PARAMÉTRICO. ....     | 110 |
| Tabela 20: ANÁLISE DA COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE BRASIL E ESPANHA PELO TESTE NÃO PARAMÉTRICO ...                                 | 111 |
| Tabela 21: DESCRIÇÃO DOS DADOS RELACIONADOS À COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE BRASIL E MÉXICO PARA O TESTE NÃO PARAMÉTRICO. ....      | 111 |
| Tabela 22: ANÁLISE DA COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE BRASIL E MÉXICO PELO TESTE NÃO PARAMÉTRICO .....                                | 112 |
| Tabela 23: DESCRIÇÃO DOS DADOS RELACIONADOS À COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE OS PAÍSES, PARA O TESTE PARAMÉTRICO. ....               | 113 |
| Tabela 24: TESTE DA VARIÂNCIA ANOVA, ENTRE OS PAÍSES PARA O TESTE PARAMÉTRICO.....   | 113 |
| Tabela 25: COMPARAÇÃO MÚLTIPLA ENTRE AS MÉDIAS DE ALUNOS POR TURMA ENTRE BRASIL (1), ESPANHA (2), MÉXICO (3) E ITÁLIA (4), PELO TESTE DE TUKEY. .... | 114 |
| Tabela 26: DESCRIÇÃO DOS RANKS RELACIONADOS À COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE OS PAÍSES .....   | 115 |
| Tabela 27: ANÁLISE DE TESTE KRUSKAL-WALLIS, ANALISANDO SE PELO MENOS UMA DAS MÉDIAS DE ALUNOS POR TURMA DIFERE ENTRE OS PAÍSES .....                 | 115 |

## LISTA DE SIGLAS

|         |   |
|---------|---|
| ABE     | – Associação Brasileira de Estatística  |
| AGEE    | – Agenda Global Estruturada para a Educação                                       |
| ANA     | – Avaliação Nacional da Alfabetização   |
| ANEB    | – Avaliação Nacional da Educação Básica   |
| ANRESC  | – Avaliação Nacional do Rendimento Escolar  |
| BIAC    | – Business Industry Advisory Committee to the OECD                                |
| BIRD    | – Banco Internacional de Reconstrução e Desenvolvimento                           |
| CEMC    | – cultura educacional mundial comum   |
| CONAES  | – Comissão Nacional de Avaliação da Educação Superior                             |
| CPC     | – Conceito Preliminar de Curso  |
| ENC     | – Exame Nacional de Cursos  |
| ENADE   | – Sistema Nacional de Desempenho dos Estudantes                                   |
| ENCCEJA | – Exame Nacional de Certificação de Competência de Jovens e Adultos               |
| ENCE    | – Escola Nacional de Ciências Estatística   |
| ENEM    | – Exame Nacional do Ensino Médio  |
| GEPEC   | – Grupo de Estudos e Pesquisas em Educação Comparada de Políticas Educacionais    |
| IASE    | – Association for Statistical Education   |
| IBGE    | – Instituto Brasileiro de Geografia e Estatística                                 |
| ICMC    | – Bacharelado em Estatística do Instituto de Ciências Matemáticas e de Computação |
| ICOTS   | – International Conference on Teaching Statistics                                 |
| IDEB    | – Índice de Desenvolvimento da Educação Básica                                    |
| IDEG    | – Instrument Development Expert Group   |
| IEA     | – Evaluation of Educational Achievement   |
| IES     | – Instituições de Ensino Superior   |
| INEE    | – Instituto Nacional para la Evaluación de la Educación                           |
| INEP    | – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio                   |

Teixeira

|        |   |
|--------|---|
| INES   | – Indicators of Education Systems                                   |
| ISI    | – International Statistic Institute                                 |
| MEC    | – Ministério da Educação  |
| MECD   | – Ministerio de Educación, Cultura y Deporte                        |
| NDMs   | – National Data Managers  |
| NPMs   | – National Project Managers   |
| OCDE   | – Organização de Cooperação e Desenvolvimento Econômico             |
| ONU    | – Organização das Nações Unidas                                     |
| PAIUB  | – Programa de Avaliação Institucional das Universidades Brasileiras |
| PCN    | – Parâmetros Curriculares Nacionais                                 |
| PDE    | – Plano de Desenvolvimento da Educação                              |
| PIAAC  | – Programme for the International Assessment of Adult Competencies  |
| PIRLS  | – Progress in International Reading Literacy Study                  |
| PISA   | – Programa Internacional de Avaliação de Estudantes                 |
| PROUNI | – Programa Universidade para Todos                                  |
| SAEB   | – Sistema Nacional de Avaliação da Educação Básica                  |
| SAEP   | – Sistema Nacional de Avaliação do Ensino Público de 1º grau        |
| SEIE   | – Sistema Estatal de Indicadores de la Educación                    |
| SINAES | – Sistema Nacional de Avaliação de Desempenho Superior              |
| UERJ   | – Universidade Estadual do Rio de Janeiro                           |
| UNESCO | – United Nations Educational, Scientific and Cultural Organization  |
| USP    | – Universidade Estadual de São Paulo                                |
| TALIS  | – Pesquisa Internacional sobre Ensino e Aprendizagem                |
| TIMSS  | – Trends in International Mathematics and Science Study             |
| TUAC   | – Trade Union Advisory Committee to the OECD                        |

## SUMÁRIO

|  |            |
|--|------------|
| <b>INTRODUÇÃO .....</b>  | <b>14</b>  |
| <b>1. PESQUISA EDUCACIONAL E ESTATÍSTICA.....</b>  | <b>20</b>  |
| 1.1. CONCEPÇÃO DE AVALIAÇÃO .....  | 21         |
| 1.2. AS PESQUISAS DE LARGA ESCALA E POLÍTICAS EDUCACIONAIS .....                                 | 25         |
| 1.3. INDICADORES EDUCACIONAIS E ORGANISMOS MULTILATERAIS .....                                   | 35         |
| 1.3. PESQUISA TALIS: UM EXEMPLO DE PESQUISA DE LARGA ESCALA.....                                 | 39         |
| 1.3.1 Contexto e objetivos da pesquisa.....  | 39         |
| 1.3.2. Metodologia utilizada na pesquisa TALIS.....  | 44         |
| 1.3.3 Questionários da Pesquisa TALIS.....   | 48         |
| <b>2. FUNDAMENTOS TEÓRICOS DA ESTATÍSTICA .....</b>  | <b>51</b>  |
| 2.1 A ESTATÍSTICA NA HISTORIA .....  | 51         |
| 2.2. EDUCAÇÃO E ESTATÍSTICA.....   | 55         |
| 2.3. POPULAÇÃO E AMOSTRA .....   | 58         |
| 2.4. ESTATÍSTICA, CONCEITOS E MÉTODOS.....   | 62         |
| 2.5. A ESTATÍSTICA DESCRITIVA .....  | 64         |
| 2.5.1 Conceito e Finalidades .....   | 65         |
| 2.5.2. Escalas de Medida .....   | 67         |
| 2.5.3. Medidas de Tendência Central e Dispersão.....   | 70         |
| 2.6. ESTATÍSTICA INFERENCIAL.....  | 78         |
| <b>3 ANÁLISE DA ESTATÍSTICA NAS PESQUISAS EDUCACIONAIS .....</b>                                 | <b>85</b>  |
| 3.1. ANÁLISE DA ELABORAÇÃO DOS QUESTIONÁRIOS E DAS ESCALAS<br>UTILIZADAS .....                   | 85         |
| 3.2. ANÁLISE SOBRE USO DO DESVIO E DO ERRO PADRÃO .....  | 100        |
| 3.3. COMPARAÇÃO DE MEDIDAS DE TENDÊNCIA CENTRAL .....  | 104        |
| 3.3.1. Comparação com base em testes paramétricos entre duas amostras<br>independentes .....     | 104        |
| 3.3.2. Comparação com base em testes não paramétricos entre duas amostras<br>independentes ..... | 110        |
| 3.3.3. Comparação entre três ou mais amostras independentes .....                                | 112        |
| <b>Considerações Finais .....</b>  | <b>117</b> |
| <b>REFERÊNCIAS.....</b>  | <b>124</b> |



## INTRODUÇÃO

Esse trabalho tem como objeto o uso da estatística como ferramenta utilizada nas pesquisas educacionais, tomando como foco central o emprego de medidas de tendência central e de dispersão. Para avaliar o emprego adequado dessas medidas, toma-se como exemplo a *Pesquisa Internacional sobre Ensino e Aprendizagem* (TALIS)<sup>1</sup>, realizada pela *Organização de Cooperação e Desenvolvimento Econômico* (OCDE), por meio do banco de dados disponibilizados por esta organização e dos dados apresentados no relatório publicado em 2009.

A definição do tema em questão nasceu das necessidades geradas pela pesquisa *O aporte das pesquisas internacionais de avaliação do ensino e da aprendizagem na educação brasileira e as políticas públicas de formação e trabalho docente*, desenvolvida pelo Grupo de Estudos e Pesquisas em Educação Comparada de Políticas Educacionais (GEPEC). Naquela investigação foram selecionados os casos do Brasil, da Espanha e do México para realizar estudos comparados a partir da pesquisa TALIS e do Programa Internacional de Avaliação de Estudantes (PISA), ambos desenvolvidos pela OCDE.

O que este estudo busca é demonstrar, por meio da aplicação das medidas de tendência central e de dispersão na pesquisa TALIS, a necessidade de averiguar a adequação ou inadequação das estatísticas utilizadas em pesquisas de larga escala, em âmbito mundial e nacional, e os impactos que poderão acarretar nas tomadas de decisão relacionadas às políticas educacionais.

A relevância desta investigação justifica-se pelo fato de a estatística se encontrar cada vez mais presente no âmbito da pesquisa em políticas educacionais. Estudar estas questões implica em entender o emprego da estatística no campo da educação, com a finalidade de medir a qualidade dos sistemas educativos e seus desempenhos.

A pesquisa TALIS, tomada como base para o estudo pretendido, consiste em uma pesquisa de alcance mundial, desenvolvida pela OCDE juntamente com um grupo de consórcio, com o objetivo de levantar dados sobre professores e diretores das escolas públicas e privadas. Assim:

---

<sup>1</sup>(LivreTradução do autor). Do original: *Teaching and Learning International Survey*.

TALIS surge da revisão da OCDE de 2005 sobre as políticas dos professores, quando identificou lacunas importantes nos dados internacionais e tem como objetivo ajudar os países a revisar e desenvolver políticas para tornar a profissão de professor mais atrativa e efetiva. TALIS é um programa de pesquisas, desenvolvido em 'rodadas' para levantar problemas elencados por cada um dos países envolvidos<sup>2</sup>(OCDE, 2009, p.03).

O Brasil participa dessa pesquisa, coordenada internacionalmente pela OCDE, e, em âmbito nacional, pelo *Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira* (INEP). No ano de 2007-08, foi realizada a primeira rodada da pesquisa, cujo foco principal foi o ambiente de aprendizagem e as condições de trabalho que as escolas oferecem aos professores dos anos finais do ensino fundamental. Em 2012-13, o Brasil e mais 31 países fizeram parte da segunda rodada (INEP, 2011), que não será analisada neste trabalho, considerando que o relatório foi publicado em 2014, com exceção da comparação entre os questionários das duas etapas para averiguar mudanças entre ambos.

Recentemente, a melhoria da qualidade educacional tem-se constituído em um objetivo político crucial na maioria dos países e, em grande medida, passou a ser também assumido no Brasil, especialmente a partir das reformas educacionais implementadas, notadamente, após a década de 1990 (CASASSUS, 2001; FERREIRA e OLIVEIRA, 2009).

Segundo a OCDE, a finalidade da pesquisa TALIS é orientar a implementação de políticas relacionadas com a formação e a condição de trabalho de professores, pelos países que fazem parte deste inquérito. Para isso, a organização oferece um banco de dados de considerável alcance (OCDE, 2009), que foi útil para os estudos aqui apresentados.

Porém, os dados proporcionados pelo relatório da OCDE limitam-se à utilização de técnicas de estatística descritiva, cujas análises são baseadas apenas em médias e percentuais. Nesse caso, é importante destacar a utilização das medidas de dispersão, ou seja, o desvio padrão e o coeficiente de variação; porque pensar em uma análise descritiva somente com a média propriamente dita não é suficiente para dar a conhecer uma realidade por meio de uma pesquisa. Em

---

<sup>2</sup> (LivreTradução do autor). No original: *TALIS draws on the OECD's 2005 review of teacher policy, which identified important gaps in international data, and aims to help countries review and develop policies to make the teaching profession more attractive and more effective. TALIS is conceptualised as a programme of surveys, with successive rounds designed to address policy-relevant issues chosen by countries.*

síntese, precisa-se conhecer a dispersão, para depois decidir sobre qual medida deve ser tomada para análise: a média, se esta for adequada ou, caso contrário, a mediana ou a moda, que podem apresentar um uso mais confiável no campo da educação.

Um exemplo muito comum sobre o uso da média é apresentado nos resultados da pesquisa TALIS, no qual o Brasil exibe uma média de 32 alunos por turma, México 37 e a Espanha 22 (OCDE, 2009). Ora, há de se pensar que, se a média destes países está em 32, 37 e 22, respectivamente, existem, por trás desses números, escolas com maior e menor número de alunos por turma, fato que pode ou não interferir na média. Partindo desse pressuposto, esse tipo de levantamento de dados será tomado como base, para analisar o emprego da média e da mediana, identificando as medidas de dispersão, como forma de demonstrar a viabilidade ou não do uso dos seus resultados em termos de estatística descritiva e aplicar a estatística inferencial para concluir a análise em questão.

O estudo da escolha e da aplicação dos métodos estatísticos possibilita uma análise mais apurada da realidade escolar e das políticas educacionais que a determina. Considerando que a Estatística pode influenciar as políticas educacionais:

[...] num mundo em constante mudança, a planificação e a tomada de decisões deverá apoiar-se em bases sólidas, no conhecimento profundo das situações passadas e presentes e numa previsão fundamentada da evolução futura (REIS, 2008, p. 16).

Assim, um estudo sobre o uso da estatística no inquérito TALIS permite compreender melhor os limites e as possibilidades desse tipo de pesquisa, que tem como finalidade avaliar e orientar as políticas educacionais que norteiam o sistema de ensino de cada país.

Assim, indica-se como objetivo geral:

—Avaliar o uso da Estatística, com vistas a subsidiar a avaliação e implementação de políticas educacionais, tomando como referência a pesquisa TALIS.

Como objetivos específicos, propõem-se:

—Analisar a importância das estatísticas utilizadas em pesquisas de larga escala para avaliação de políticas educacionais;

- Situar a avaliação como um processo de análise e acompanhamento da implementação de políticas, cujos indicadores são elaborados a partir de procedimentos estatísticos;
- Identificar os limites das pesquisas utilizadas como instrumento de controle no processo de avaliação externa, pelos governos e organismos multilaterais;
- Compreender os conceitos, finalidades e métodos estatísticos para subsidiar o uso da Estatística nas pesquisas educacionais;
- Avaliar a importância da elaboração dos questionários para a pesquisa, a partir das questões formuladas na pesquisa TALIS (OCDE);
- Verificar a adequação da aplicação das medidas de tendência central e de dispersão, a partir do banco de dados da pesquisa TALIS, a fim de compreender a realidade investigada;
- Realizar a comparação de médias, por meios de testes inferências, para possibilitar uma conclusão mais precisa, ultrapassando a mera descrição.

Esta investigação toma como princípio norteador a análise da aplicação dos métodos estatísticos como instrumento para explicar a realidade educacional, com a pretensão de permitir a comparação entre grupos (países, estados, municípios, escolas, etc.). Pesquisas desse tipo classificam países, escolas, professores e estudantes, a partir da ótica dos seus promotores. E, assim:

Quando se tornam públicos os resultados da avaliação dos alunos, das escolas e dos profissionais da educação, identificam-se as médias mais altas e mais baixas em um sistema de classificação que não permite conhecer os múltiplos fatores que influíram em todo o processo educativo e, muito menos, no próprio processo de avaliação.

Nesse sentido, é fundamental considerar, antes de tudo, aspectos de ordem teórica e metodológicos que não ignorem os múltiplos fatores que exercem influência sobre o objeto de comparação, para não chegar a resultados falsos ou equivocados (TROJAN; SÁNCHEZ, 2009, p.2).

Desse modo, deve-se ter atenção com a questão metodológica, considerando que envolvem princípios de ordem epistemológica e pressupostos teóricos que refletem o ponto de vista de acordo com o qual se interpreta um determinado objeto.

Assim, os procedimentos metodológicos devem tomar como ponto de partida um estudo teórico sobre as finalidades de pesquisas sobre avaliação de políticas educacionais que utilizam dados numéricos para compreender e avaliar a realidade da educação, bem como os fundamentos, conceitos e métodos da Estatística, que se coloca como uma das principais ferramentas.

Segundo Azevedo e Campos (1981) a Estatística como ciência apresenta um conjunto de passos denominados fases do método estatístico, os quais se identificam com o método de pesquisa em geral: *definição do problema*, que consiste na definição e formulação correta do problema a ser estudado; *planificação*, que consiste em determinar o processo, como obter informações sobre a variável ou variáveis em estudo; *coleta dos dados*, que pode ser obtida por meio de questionários, observação, experimentação, ou pesquisa bibliográfica; *organização dos dados*, que consiste em resumir os dados através de sua contagem e agrupamento. Finalmente, obtém-se um conjunto de informações que irá conduzir o estudo estatístico, do qual deriva a *análise e interpretação dos dados* que incluem o cálculo de novos números, com base nos dados estatísticos.

Assim, para a realização desse trabalho, o primeiro passo Foi situar a estatística no campo da pesquisa para avaliação de políticas educacionais, no atual contexto de influência da globalização política e econômica. Para isso, é fundamental o estudo teórico sobre estatística descritiva e inferencial, a fim de destacar aspectos e técnicas mais importantes para uso na pesquisa educacional.

Assim, esta investigação, inserida no campo dos estudos em políticas educacionais, se propõe a analisar alguns elementos relevantes, do ponto de vista da aplicação da estatística. Para tanto, foram analisados os questionários, as escalas de medida, as medidas de tendência central e dispersão e comparações de média, na pesquisa TALIS, conforme já mencionado. Para isso, foram utilizados os programas estatísticos JMP 8.0 e SPSS 13.0.

A exposição deste trabalho está organizada em 3 capítulos. O primeiro, denominado *Pesquisa Educacional e Estatística*, tem a finalidade de conceituar avaliação como procedimento importante adotado a partir das reformas de Estado dos anos 1990 e analisar a aplicação da estatística nas pesquisas educacionais de larga escala, especialmente aquelas adotadas pelos governos nacionais e organismos multilaterais, para avaliar a qualidade da educação por meio de indicadores, formulados a partir da Estatística. Ainda neste capítulo, descreve-se a Pesquisa TALIS – contexto e objetivos, metodologia utilizada e questionários – para tomá-la como exemplo de pesquisa de larga escala, nas análises realizadas neste estudo.

No segundo capítulo, *Fundamentos Teóricos da Estatística*, aborda-se a Estatística como ferramenta de pesquisa, a partir de sua história como ciência, como profissão e constituição dos cursos de graduação. Em seguida, são apresentados os conceitos, finalidades, os métodos da Estatística; assim como a estatística descritiva, suas escalas de medidas, medidas de tendência central e dispersão e a estatística inferencial, com seus conceitos e métodos.

Finalmente, no último capítulo, *Análise da Estatística nas Pesquisas Educacionais*, é realizada uma análise da elaboração dos questionários e das escalas; uma análise sobre o uso do desvio e do erro padrão; e a comparação de medidas de tendência central com base em testes paramétricos e não paramétricos entre duas amostras independentes e, também entre três ou mais amostras independentes, através da questão referente à média de alunos por turma, a partir do relatório e do banco de dados da OCDE.

A partir deste estudo, serão identificados os entraves na formulação das questões dos questionários, que pouco utilizam as escalas intervalares e de razão, e são apresentadas, na maior parte, sob a forma de variáveis nominais e ordinais – como a Escala Likert, por exemplo –, que limitam a confiabilidade da análise e interpretação dos dados.

## 1. PESQUISA EDUCACIONAL E ESTATÍSTICA

“Parece não haver atividade humana cujo valor não possa ser melhorado injetando-se ideias estatísticas no planejamento e usando-se métodos estatísticos para análise eficiente de dados e acesso dos resultados para realimentação e controle.” (Calyampudi Radhakrishna Rao, Estatístico indiano).

A estatística é empregada largamente em todas as áreas, como também pode ser aplicada à educação de forma satisfatória, na medida em que lida com quantidades e qualidades para o estabelecimento de padrões. Desse modo, contribui para a avaliação e formulação de políticas educacionais, tal como se observa na afirmação de Medeiros, quando assevera que os dados quantitativos podem ser úteis, ou seja:

A despeito do que possa ser considerada grande quantidade, não restam dúvidas à sua fértil aplicação no campo educacional, como ferramenta para a formulação de planos, programas e projetos nos sistemas de ensino, bem como, no interior da própria escola (MEDEIROS, 2007 p.20).

Assim sendo, o ponto de partida deste estudo é, conceituar avaliação e verificar o papel que a avaliação externa tem assumido neste contexto, com base em Souza (2009), Freitas (2005), Brito (2008) e Dias Sobrinho (2010). Em seguida, considerando as reformas de Estado ocorridas nos anos 1990, quando se estabelecem os sistemas nacionais de avaliação, discutem-se as influências das pesquisas de larga escala internacional nas políticas educacionais. Para tanto, Cassasus (2001), Afonso (2001) e Souza (2002) apresentamos fundamentos destas reformas, que tem a avaliação como um dos seus elementos centrais.

Em seguida, são consideradas as pesquisas de larga escala para, avaliação de políticas educacionais, tomando como principais referências Freitas (2007), Vianna (2003) e Minhoto (2011). A partir desse ponto de vista, evidencia-se o uso da estatística e seus consequentes indicadores educacionais utilizados por organismos multilaterais, como a OCDE, responsável pela realização da pesquisa TALIS, que é tomada neste estudo como exemplo para analisar a utilização da estatística.

## 1.1. CONCEPÇÃO DE AVALIAÇÃO

A avaliação é uma etapa importante no processo de desenvolvimento em diversas áreas: indústria, comércio, prestação de serviço, educação, etc. É através da avaliação que se podem tomar decisões relacionadas às mudanças e melhorias em uma empresa, uma indústria ou em uma escola. Assim, no âmbito da educação:

Utilizando-se dos termos mais comuns para definir avaliação, teríamos esta como um processo de análise e acompanhamento da implementação de políticas, subsidiando seu planejamento desde o diagnóstico até seus resultados, buscando orientar as tomadas de decisões (SOUZA, 2009, p. 22).

Desde os anos de 1950, a pesquisa, a avaliação, a estatística, a informação e o planejamento foram reconhecidos como recursos importantes para que os países pudessem governar a educação básica em seus territórios. A presença da avaliação educacional apareceu como uma recomendação aos países em desenvolvimento na obra da UNESCO *L'éducation dans le monde* volume V: *Politique, législation et administration de l'éducation*, publicada entre 1954 e 1959, indicando recomendações para os Estados-membros, condições necessárias ao planejamento educacional e busca de métodos mais aperfeiçoados para medir os resultados do processo educacional, dando início à construção da padronização de testes. Essa obra também ressaltou a importância da avaliação de larga escala (FREITAS, 2005).

Sobre a importância do uso sistemático da avaliação dos sistemas de ensino na época, Freitas, a define como:

[um] potencial para obter informações precisas sobre os problemas educacionais. Isso porque a falta de tais informações passava a ser vista como fator a impedir o planejamento de longo prazo nos países em desenvolvimento. Assim, a avaliação da eficácia dos sistemas nacionais de ensino, pelos governos, era ali enfatizada e recomendada aos Estados-membros como condição necessária ao planejamento educacional (FREITAS, 2005 p. 83).

Em se tratando de educação, Dias Sobrinho (2010) argumenta que a avaliação possui um papel importante na organização, criação e implementação das reformas educacionais, tendo como finalidade promover mudanças nos currículos, nas metodologias do ensino, na gestão, nos modelos institucionais e motivar a responsabilidade social. A respeito de avaliação educacional, Klein e Fontanive a definem como:



um sistema de informações que tem como objetivos fornecer diagnóstico e subsídios para a implementação ou manutenção de políticas educacionais. Ela deve ser concebida também para prover um contínuo monitoramento do sistema educacional com vistas a detectar os efeitos positivos ou negativos de políticas adotadas (KLEIN; FONTANIVE, 1995 p. 29).

De um modo geral, classificam-se os estudos avaliativos como quantitativos ou qualitativos. Entretanto, esta polêmica entre qualidade e quantidade não será abordada nesta investigação, dado que: “Quantidade e qualidade são partes integrantes e indissociáveis para se definir critérios de avaliação da qualidade” (SOUZA, 2009, p.22). O enfoque quantitativo, que tem sido utilizado pelas pesquisas de larga escala são, em geral, estudos marcados pela estatística descritiva, que não é conclusiva.

Para que seja efetiva, a avaliação não pode ser realizada de qualquer maneira. É necessário avaliar um determinado momento – que corresponde ao final de um processo – mas, também o movimento histórico de desenvolvimento dos sistemas. Neste sentido, Brito (2008) conceitua dois diferentes tipos de avaliação: *estática* e *dinâmica*. A diferença entre estática e dinâmica corresponde ao próprio significado dos termos. A primeira refere-se à avaliação de um momento específico, a partir de um instrumento padronizado e a segunda, pretende medir um processo que se desenvolve em um determinado período de tempo e inclui estudo do contexto onde se realiza e vários instrumentos de medida.

Segundo Brito (2008), no caso da avaliação do desempenho dos alunos, pode-se afirmar que a avaliação estática é composta de uma prova de estilo tradicional, que os alunos respondem de forma direta, ou seja, somente com o conhecimento aprendido ou estudado para determinada avaliação, sem *feedback*. O resultado final é, quase sempre, “a soma da pontuação dos itens resolvidos corretamente” e, nesse sentido, é comum o “desconto de pontos (penalidade) nas questões em cuja resposta o aluno não tem segurança (chute)” (BRITO 2008 p. 845). Exemplos desse tipo de avaliação são os concursos, vestibulares, pesquisas de larga escala, relacionadas ao desempenho dos alunos (Prova Brasil, SAEB, PISA, etc.), aos professores ou aos sistemas educacionais (Pesquisa TALIS).

Na avaliação dinâmica, Brito (2008 p.845) informa que os alunos aprendem “durante o intervalo entre as testagens e com as provas”. Busca avaliar o desenvolvimento do aluno em relação aos novos conhecimentos e como estes são

trabalhados, é possível dar o *feedback*, auxiliá-lo na resolução de problemas e o interesse principal é “no processo, pois, em um mundo em constante transformação, o conhecimento deve ser uma ferramenta útil para o estudante, ele deve tornar-se capaz de dominar e transferir o aprendido para novas situações”.

Neste sentido, pode-se afirmar que nas pesquisas de larga escala, realizadas em âmbito nacional e internacional, não cabe uma avaliação dinâmica, tendo em vista que estas não incluem estudo sobre os processos e os contextos relacionados com o objeto da avaliação.

No caso da pesquisa TALIS, pode-se afirmar que a análise realizada pelo relatório (OCDE, 2009) apresenta uma avaliação estática sobre a formação e as condições de trabalho dos professores investigados, pois se limita à descrição dos resultados obtidos por apenas um instrumento. Exemplo: Todo relatório da pesquisa TALIS apresenta tabelas descritivas levantadas através de um questionário para os professores e outro para os diretores, “preparados para explorar questões analíticas e políticas sugeridas pelos países participantes sob os termos consentidos” (OCDE, 2009 p. 16).

Mesmo considerando a segunda rodada da pesquisa realizada em 2012-13, a comparação dos resultados das duas aplicações não pode ser considerada como uma avaliação dinâmica. Pode-se apenas identificar as mudanças ocorridas entre os dois períodos de aplicação e incluir outros instrumentos que investiguem os porquês das mudanças, além de realizar a aplicação de estatística inferencial, que permita analisar e comparar países, como por exemplo, porte das escolas, número de alunos por sala, etc.

Entretanto, o questionário aplicado em 2012-13, apresentou alterações que, de um lado, permitiram o aperfeiçoamento do instrumento e correção de alguns equívocos – o que é desejável. Mas, por outro lado, causa problemas de comparação porque as respostas foram elaboradas com outro tipo de escala, dificultando a comparação dos questionários. Exemplo: No questionário de 2007-8, a questão nº 2, para os professores – Qual sua idade? – apresenta opções de resposta em escala ordinal e no questionário de 2012-13, a mesma questão está em escala de razão, ou seja, utilizam-se escalas diferentes, que serão explicadas mais adiante. Como houve uma alteração, ao confrontar as questões, não se podem desenvolver análises sobre as mudanças, somente descrever.

Para avaliar políticas educacionais é fundamental utilizar-se uma avaliação dinâmica, ou seja, partir de diagnóstico e da implementação das políticas, acompanhar o processo, incluir medidas corretivas, para avaliar os resultados obtidos e a sua eficácia. Os sistemas de avaliação de larga escala – como, por exemplo, a Prova Brasil, o Saeb, o PISA, etc. – utilizam o teste com único instrumento de avaliação, ou seja, de forma estática. Os resultados de tais instrumentos são avaliados entre si, portanto, é feita apenas a comparação de um ano específico com outro posterior. Para uma avaliação dinâmica faz-se necessário incluir outros recursos relacionados ao contexto e ao processo. Como por exemplo, quando se comparar as mudanças ocorridas na formação inicial dos docentes a partir das duas aplicações da pesquisa TALIS, devem-se analisar também as políticas adotadas, a legislação vigente, os problemas ocorridos no período, a relação entre oferta e demanda de vagas docentes, os planos de cargos e salários, os currículos dos cursos, entre outros fatores, que ultrapassam os limites deste tipo de pesquisa.

Porém, a avaliação estática, que apresenta dados numéricos pode ser útil como referência para pesquisas relacionadas à avaliação de políticas, no sentido de dimensionar as condições oferecidas para realização do processo educacional que se avalia. Entretanto, estudos que utilizam mensuração de dados são poucos e esses, na sua maior parte, são apenas descritivos. GATTI (2004) introduz um questionamento contundente sobre a questão:

Essa dificuldade no uso de dados numéricos na pesquisa educacional rebate de outro lado na dificuldade de leitura crítica, consciente, dos trabalhos que os utilizam, o que gera na área educacional dois comportamentos típicos: ou se acredita piamente em qualquer dado citado (muitas vezes dependendo de quem cita — argumento de autoridade), ou se rejeita qualquer dado traduzido em número por razões ideológicas reificadas, *a priori* (GATTI, 2004 p.13).

Cabe lembrar também que, cada tipo de dado implica em certos tipos de tratamento estatístico possíveis e, para isso, ressalta-se aqui a importância do uso das medidas apropriadas. GATTI acrescenta que:

... as medidas [...] não podem ser tomadas como sendo a própria natureza das coisas; isto também se aplica às categorias dos estudos de análise de conteúdo e outras análises dos modelos qualitativos. O que se procura ao criar uma tradução numérica ou categoria de fatos, eventos, fenômenos, é que esta tradução tenha algum grau de validade

racional, teórica, no confronto com a dinâmica observável dos fenômenos (GATTI, 2004 p.15).

A discussão sobre a qualidade dos dados estatísticos – como os censos, pesquisas de larga escala por amostragem, entre outros – vem de longos anos. Muitos autores se debruçam sobre esse assunto para analisar os limites e as possibilidades das pesquisas. Por mais que existam questionamentos sobre a natureza dos dados, é evidente que sem esses dados, muitas questões de natureza sócio educacionais não poderiam ser dimensionadas, equacionadas e compreendidas e algumas nem seriam levantadas (GATTI, 2004).

## 1.2. AS PESQUISAS DE LARGA ESCALA E POLÍTICAS EDUCACIONAIS

Qualquer problema induz ao levantamento de hipóteses e, de alguma forma, a resolução desses implica em entender o que está a sua volta. Aqui, analisa-se a utilização da Estatística em pesquisas de larga escala para entender como esta ciência influencia o campo da avaliação de políticas educacionais, tornando-se um veículo essencial para tomadas de decisão, e, ao mesmo tempo, mostrando que, quando empregada erroneamente, a estatística pode levar a resultados desastrosos.

No caso, a estatística é apenas um instrumento a serviço da educação e existem duas atitudes contraditórias e igualmente errôneas do pesquisador em relação às suas técnicas. A primeira é pensar que a estatística é sempre capaz de fornecer resultados conclusivos através de seus resultados obtidos, como é o caso das pesquisas realizadas por organizações internacionais e órgãos oficiais de governo. A segunda seria rejeitá-la, considerando que não tem valor porque não é capaz de confirmar hipóteses, como é o caso de pesquisadores que somente confiam em estudos de caso. A estatística não cria informação que não existe. Seu grande mérito é colocar em evidência fatos novos, assim como melhorar os mecanismos já aplicados, destacando o que realmente é essencial em uma forma que facilite as tomadas de decisão (VALENTIM, 2000).

Para compreender o papel das pesquisas de larga escala na avaliação das políticas educacionais e a importância que lhes é atribuída, é preciso contextualizar o atual momento de globalização e as reformas de Estado, levadas a cabo a partir dos anos 1990, na maioria dos países.

Na contemporaneidade, entende-se globalização como o processo que implicou numa transformação de padrões de ordem política, econômica, social e cultural. A partir de uma nova crise do capital que teve como consequências uma reestruturação produtiva, uma nova divisão internacional do trabalho e uma expressiva expansão de novas tecnologias em todas as áreas relacionadas com a produção e com os meios de informação e comunicação, os países expandiram suas relações transnacionais, gerando um novo paradigma econômico e político de âmbito mundial. CASASSUS (2001, p. 8) comenta que uma forma de perceber esse processo é através da “desterritorialidade, na qual o Estado nacional perde suas fronteiras e emergem outros espaços para a configuração da identidade e da cultura”. Roger Dale afirma que:

[...] La globalización representa un cambio nuevo y característico de la relación entre el estado y las fuerzas supranacionales, que ha afectado a la educación de un modo profundo y de una serie de formas. Sus efectos han sido mucho más importantes y directos en los países en vías de desarrollos, antes que en los Estados dominantes, aunque ninguno ha sido inmune a ellos. [...] Los efectos de la globalización también se pueden experimentar directamente, tanto a nivel sectorial (educativo) como organizativo. [...] Una de las principales características de la globalización que presenta a todos los países una agenda mucho más común de problemas de lo que había ocurrido hasta ahora; (DALE, 2002, p.71-72).

No mesmo sentido, Freitas (2007, p.136) complementa, afirmando que a globalização implica numa transformação dos padrões dominantes de organização socioeconômica, assim como também possibilita novas formas transnacionais, que tornam as comunidades mais vulneráveis aos acontecimentos globais.

Nesse novo contexto, a mudança da economia capitalista mundial dá direção para o processo de globalização. Segundo Dale, a natureza dessa mudança deve ser investigada para identificar os seus efeitos sobre os sistemas educativos, ainda que sua influência esteja intensamente mediada pelo conjunto de instituições e modos de organização e gestão dos Estados nacionais. Dale (2004), afirma que de um lado, há uma explicação do ponto de vista de uma *agenda global estruturada para a educação* (AGEE) que exerce influência, mediada pelas forças locais, com base em estudos de Cox (1996), Mittelman (1996) e Hettne (1996); e de outro, de uma cultura educacional mundial comum (CEMC), com base nos estudos de John Meyer e outros (1992).

Assim, com a atual globalização, os países intensificaram sua preocupação com a definição de padrões de qualidade internacional para a educação. De tal modo, passam a estabelecer acordos e agendas em comum. Por intermédio de organizações multilaterais, os países, tanto os mais desenvolvidos como aqueles em desenvolvimento, passam a analisar a educação, a partir da referida AGEE (DALE, 2002). Nesse sentido, elaboram programas de pesquisa nacionais e internacionais, para analisar os sistemas de ensino, a partir das necessidades da economia.

No intuito de se modernizar e acompanhar o novo paradigma global, os países periféricos – adequados à racionalidade econômica em meio às Reformas do Estado dos anos de 1990 – começaram a adotar os mesmos padrões e critérios de produtividade econômica para acompanhar os países desenvolvidos, seguindo, como pontua SOUZA (2002), quatro grandes objetivos que podem ser definidos como centrais nas orientações internacionais:

a) melhoria da *eficácia* da atividade administrativa; b) melhoria da *qualidade* na prestação dos serviços públicos; c) diminuição das *despesas públicas*; d) aumento da *produtividade na administração* do Estado (SOUZA 2002, p. 90).

No campo educacional, os objetivos gerais das políticas adotadas seguem as estratégias e prioridades dos organismos multilaterais, particularmente dos bancos financiadores. O Banco Internacional de Reconstrução e Desenvolvimento<sup>3</sup> (BIRD), por exemplo, foca “a promoção da equalização social por meio da promoção da igualdade de oportunidades nos diferentes níveis e setores sócioeducacionais” (SOUZA, 2002, p. 90). No âmbito das reformas de Estado, a educação básica passou a ser vista como alvo principal para o desenvolvimento social e econômico dos países subdesenvolvidos, não obstante, para os Bancos multilaterais, o nível de desenvolvimento intelectual e humano está diretamente atrelado ao nível de

---

<sup>3</sup>Banco Internacional de Reconstrução e Desenvolvimento – BIRD – Em *inglês Bank for Reconstruction and Development*, (IRDB) criado em 1944 através dos acordos de *Bretton Woods* e possui atualmente 187 países-membros. Com finalidade de promover o desenvolvimento econômico e social, e a redução da pobreza, contribuindo para a melhoria da qualidade de vida das pessoas. Além de financiar projetos claramente ligados a questões acerca de infraestrutura, educação e meio ambiente, que visem acelerar o crescimento econômico brasileiro, o Banco Mundial produz estudos e relatórios diversos, além de análises adicionais que incluem revisões sobre política econômica, meio ambiente e gastos fiscais. São fornecidos dois tipos de empréstimos: a) para investimento em obras, bens e serviços, que visem apoiar projetos de desenvolvimento social e econômico, abarcando uma ampla gama de setores; b) para ajuste, utilizados para assistir às reformas institucionais e políticas do país. (<http://www.ambiente.sp.gov.br/wp-content/uploads/2011/10/bird.pdf>).

desenvolvimento social e econômico de cada país em desenvolvimento (SOUZA, 2002).

Os efeitos da globalização na educação são explicados por Dale, do ponto de vista da CEMC, pela existência de uma hipotética cultura mundial e do ponto de vista da AGEE, como uma nova forma de força supranacional que afeta os sistemas educativos nacionais. Por meio da AGEE, pode-se identificar “como os Estados interpretam e respondem a uma agenda comum” (DALE, 2004, p. 454).

Segundo Dale (2002), as organizações multilaterais passam a se constituir como o principal canal de influência. Estes organismos, com o aval dos Estados nacionais, definem as regras para as políticas locais, tanto as relacionadas com o processo de ensino e aprendizagem quanto aos sistemas de avaliação externos – nacionais e internacionais.

O contexto social e o papel da escola e do docente estão mudando e, em face desse panorama, o desenvolvimento profissional do professor é estabelecido em novas bases. Independentemente da formação, recebida antes de iniciar a profissão, novos desafios são apresentados aos professores e os sistemas de educação se responsabilizam pelo desenvolvimento profissional no decorrer da carreira do professor, para atender às necessidades e problemas que se apresentam a partir da AGEE (DALE, 2002).

Os sistemas de avaliação, neste sentido, são criados para avaliar a adequação ao modelo estabelecido. Assim, as pesquisas de larga escala são criadas para avaliar as condições dos sistemas educacionais, professores, diretores e alunos, com a intenção de criar ou melhorar políticas adotadas. Segundo Alves (2011), a origem desse tipo de avaliação remonta à década de 1950:

[...] com o desenvolvimento de métodos estatísticos, o aperfeiçoamento da sistematização da coleta de dados, a sofisticação das técnicas de pesquisa e o desenvolvimento de computadores com capacidade de processamento de grande quantidade de dados, que a avaliação de políticas públicas teve grande impulso tanto nos países desenvolvidos quanto nos países em desenvolvimento (ALVES, 2011, p. 34).

Deste modo, as avaliações nacionais passaram por uma série de modificações ao longo do século passado, até chegar ao contexto atual, caracterizando-se como principal instrumento para avaliar o sistema público de ensino. Sass e Minhoto indicam que o status da avaliação educacional no Brasil:

[...] pode ser caracterizado sumariamente como o principal instrumento de verificação da expansão, do atendimento, da eficácia e eficiência, especialmente, do sistema público de ensino, em todas as etapas da educação, além de proporcionar critérios para avaliar o desempenho escolar dos alunos, as unidades escolares quanto à qualidade do ensino e aos recursos humanos e materiais disponíveis, bem como propiciar ao aluno do ensino médio uma via de ingresso ao ensino universitário. (SASS e MINHOTO, 2011 p. 64).

O sistema de avaliação no Brasil foi se ampliando, a partir do final da década de 1980, com a criação, em 1988, do *Sistema Nacional de Avaliação do Ensino Público de 1º grau* (SAEP), um estudo piloto, cujo objetivo foi testar instrumentos e procedimentos para analisar a educação pública. Porém, a falta de recursos impediu a continuidade do projeto (SILVA, 2010).

A partir de 1990, foi criado o *Sistema Nacional de Avaliação da Educação Básica* (SAEB) com o objetivo, segundo o *Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira* (INEP) de:

Avaliar a Educação Básica brasileira e contribuir para a melhoria de sua qualidade e para a universalização do acesso à escola, oferecendo subsídios concretos para a formulação, reformulação e o monitoramento das políticas públicas voltadas para a Educação Básica. Além disso, procura também oferecer dados e indicadores que possibilitem maior compreensão dos fatores que influenciam o desempenho dos alunos nas áreas e anos avaliados (INEP, 2011).

Segundo o INEP, o Saeb é composto por três avaliações externas em larga escala: a *Avaliação Nacional da Educação Básica* (Aneb), que abrange, de maneira amostral, alunos das redes públicas e privadas do país, em áreas urbanas e rurais, do 5ºano e 9ºano do Ensino Fundamental e do 3º ano do Ensino Médio, para avaliar a qualidade, a equidade e a eficiência da educação brasileira, apresentando resultados do país como um todo, das regiões geográficas e das unidades da federação; a *Avaliação Nacional do Rendimento Escolar* (Anresc), também denominada *Prova Brasil*, que é uma avaliação censitária, envolvendo alunos do 5º ano e 9º ano do Ensino Fundamental das escolas públicas das redes municipais, estaduais e federal, para avaliar a qualidade do ensino das escolas públicas; e a *Avaliação Nacional da Alfabetização* (ANA), que também é censitária, envolvendo os alunos do 3º ano do Ensino Fundamental das escolas públicas, para avaliar os níveis de alfabetização e letramento em Língua Portuguesa, alfabetização Matemática e condições de oferta do *Ciclo de Alfabetização* das redes públicas (INEP, 2011).



Figura 1: COMPOSIÇÃO DO SAEB



Fonte: INEP, 2011.

Minhoto (2011) afirma que a o SAEB e a Prova Brasil expandiram-se para ocupar um lugar de destaque na definição das políticas do MEC. Informa ainda que, em 2007, o Governo Federal desenvolveu o *Plano de Desenvolvimento da Educação* (PDE), contendo 30 metas e ações, com o objetivo de melhorar a qualidade do ensino de cada escola e rede de ensino, monitoradas através do *Índice de Desenvolvimento da Educação Básica* (IDEB), que é composto por outros dois indicadores: taxa média de aprovação e a pontuação média obtida na Prova Brasil. O IDEB é calculado pelo produto entre o desempenho e o rendimento escolar, variando de 0 a 10. Esse índice é medido a cada dois anos e tem como meta chegar a 6 pontos até 2022. Essa meta é equivalente à média estabelecida entre os países membros da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) – fato que evidencia a influencia desta organização nas políticas educacionais.

Segundo o INEP, a fórmula do IDEB é a seguinte:  $N = n \times p$ ; onde  $n$  corresponde aos valores dos limites inferior e superior, obtidos a partir da média e desvio padrão da prova de proficiência em Matemática e Língua Portuguesa, divididos por dois; e o  $p$  é obtido diretamente do censo escolar, pelo tempo médio (em anos) para conclusão de uma série na etapa do ensino (INEP 2011). Entretanto, o IDEB é um índice que, no fim das contas, limita-se à medição das condições dos alunos, e assim:

O verdadeiro limite à universalização da melhoria da qualidade da escola é a própria ideologia meritocrática liberal. Caso a avaliação se coloque a serviço dela, então ficará limitada à medição do mérito e à ocultação da desigualdade social sob a forma de indicadores “neutros” como o Índice

de Desenvolvimento da Educação Básica (IDEB) criado pelo MEC (FREITAS, 2007 p. 971)

O *Exame Nacional do Ensino Médio* (ENEM) foi criado em 1998, com o objetivo de avaliar o desempenho do estudante ao concluir a educação básica, buscando contribuir com subsídios para a melhoria da qualidade desse nível de escolaridade. Em 2005, passou a ser utilizado como critério de seleção para bolsas de estudos no *Programa Universidade para Todos* (ProUni). A partir de 2009, o ENEM passou por uma nova reformulação e oficializou o exame como forma de seleção nos processos seletivos das universidades federais, tal como observa Minhoto:

Essas instituições de ensino tiveram autonomia para optar entre quatro possibilidades de utilização do NOVO ENEM em seus processos seletivos: como fase única; como primeira fase; combinado o exame com o vestibular da instituição; e como fase única para as vagas remanescentes do vestibular (MINHOTO 2011 p. 179).

A partir dessa reestruturação, o MEC, além de democratizar o acesso às vagas federais de ensino superior, passou a possibilitar a mobilidade acadêmica e induzir a reestruturação dos currículos do ensino médio (MINHOTO 2011).

Em 2002, foi criado o *Exame Nacional de Certificação de Competência de Jovens e Adultos* (ENCCEJA) para medir habilidades e competências na conclusão do ensino fundamental e médio. Tem como objetivo validar e certificar essa população com uma prova de proficiência em cada área do conhecimento, que varia de 60 e 180 pontos, ou seja, se o aluno conseguir a média em determinada área estará apto às exigências mínimas necessárias para obter a certificação. Minhoto afirma que:

Entretanto, o ENCCEJA parece valorizar menos as condições de acesso aos conhecimentos socialmente valorizados por jovens e adultos e mais a aquisição de uma certificação, ao reforçar a importância da obtenção do diploma (MINHOTO, 2011 p. 182).

Além disso, a vantagem das técnicas que se limitam a medir competências e habilidades e não o conhecimento, de acordo com Freitas (2007), é aparente, pois os sistemas escolares, escolas e redes de ensino diferem quanto aos conteúdos e formas de ensino. E, desse modo, as discrepâncias podem esconder diferenças de ordem curricular e pedagógica e não de aprendizagem ou aquisição de competências e habilidades.

As Políticas de Avaliação da Educação Superior vigoram a partir da década de 1990, através do *Programa de Avaliação Institucional das Universidades Brasileiras* (PAIUB), criado em 1993, e o *Exame Nacional de Cursos* (ENC), criado em 1996, conhecido como *Provão*.

A base do PAIUB era a avaliação institucional através das próprias *Instituições de Ensino Superior* (IES), com o objetivo de qualificar as diversas dimensões da vida acadêmica. Como se refere Minhoto é composta por quatro instrumentos de avaliação:

Diagnóstico; Avaliação Interna (cursos, disciplinas, docentes, discentes, pessoal técnico administrativo, infraestrutura, currículo, relação instituição e sociedade, egressos) e Reavaliação. (MINHOTO, 2011 p. 183).

Esse programa enfatizou métodos de avaliação qualitativa, respeitando a identidade de cada IES, sua adesão era voluntária e não resultava em qualquer tipo de punição, decorrente dos resultados obtidos (MINHOTO 2011). A origem do PAIUB, entretanto, é anterior a sua institucionalização. De acordo com Michelotto:

De fato, em 1982, a ANDES (Associação Nacional dos Docentes da Educação Superior) propôs uma avaliação institucional que foi encampada por algumas universidades e, após marchas e contramarchas, contando com a participação de outras associações e fóruns de docentes e também com a aceitação do governo, acabou por gerar a institucionalização do PAIUB (Programa Nacional de Avaliação Institucional das Universidades Brasileiras), em 1993 (MICHELOTTO, 1999 p. 116).

O ENC foi criado pelo governo federal, tendo como foco o exame universal e obrigatório do desempenho de todos os concluintes, em campos de conhecimentos predefinidos. Esse desempenho nas provas foi considerado como o *produto final* (output) das IES, desempenho expresso por meio das notas obtidas. O principal indicador que se destacou no ENC foi o rendimento dos concluintes, expressando uma visão estreita de avaliação do sistema de ensino, “assentada na correlação imediata entre o que é aprendido pelo aluno e aquilo que lhe é ensinado” (MINHOTO, 2011, p. 184).

Em 2004, foi criado o *Sistema Nacional de Avaliação de Desempenho Superior* (SINAES), para avaliar a responsabilidade social, o desempenho dos alunos, a gestão da instituição, o corpo docente, as instalações e vários outros aspectos, a partir de três eixos: o ensino, a pesquisa e a extensão. Possui uma série

de instrumentos complementares, tais como a autoavaliação, avaliação externa, Enade, avaliação dos cursos de graduação e instrumentos de informação (censo e cadastro) (MINHOTO 2011). Como informa o INEP:

Os resultados das avaliações possibilitam traçar um panorama da qualidade dos cursos e instituições de educação superior no País. Os processos avaliativos são coordenados e supervisionados pela Comissão Nacional de Avaliação da Educação Superior (Conaes). A operacionalização é de responsabilidade do Inep (INEP, 2011).

O Sistema Nacional de Desempenho dos Estudantes (ENADE), que integra o SINAES, tem como objetivo, de acordo com o INEP (2011), “aferir o desempenho dos estudantes em relação aos conteúdos programáticos previstos nas diretrizes curriculares do respectivo curso de graduação, e as habilidades e competências em sua formação”. O problema, segundo Minhoto, é que as informações para a elaboração do Conceito Preliminar de Curso (CPC), um dos indicadores de avaliação criado em 2008, são retiradas quase que totalmente do ENADE, sobrevalorizando este instrumento em função dos outros, que são utilizados no SINAES, “tornando-o central para a definição da qualidade da educação superior” (MINHOTO, 2011, p. 187).

Assim sendo, constata-se que a avaliação ganhou centralidade como mecanismo de controle central do Estado e,

[passou] a ser um dos eixos estruturantes das políticas públicas, em geral, e das políticas educacionais, em particular. [...] são também cada vez em maior número os estudos de análise comparada de resultados, procedimentos e políticas de avaliação que procuram ter em consideração as especificidades políticas, administrativas, ideológicas e as tradições culturais dos diferentes países (AFONSO, 2001, p. 27).

Desse modo, sistemas nacionais de avaliação foram criados em âmbito mundial. A título de exemplo, citamos os sistemas do México e da Espanha, países que são utilizados neste estudo para análise de estatísticas descritivas e inferenciais.

No México, foi criado o *Instituto Nacional para La Evaluación de La Educación* (INEE) por decreto presidencial em 8 de agosto de 2002. Nesta mesma lei, sob a responsabilidade do INEE, é constituído o *Sistema Nacional de Evaluación Educativa*, de acordo com o estabelecido na Constituição Mexicana (INEE, 2013).

Na Espanha, o *Sistema Estatal de Indicadores de La Educación* (SEIE), desde o ano 2000, coordenado pelo *Ministerio de Educación, Cultura y Deporte*

(MECD) publica uma síntese dos dados educativos nacionais e internacionais que permitem situar o país no marco da OCDE e da União Europeia. Sua elaboração se realiza conjuntamente com o *Instituto Nacional de Evaluación Educativa* (INEE). Ainda que a maioria dos indicadores tenha se mantido em todas as edições, a aparição de estudos internacionais como TIMSS, PIRLS, PISA, PIAAC y PISA Resolución de problemas<sup>4</sup>, levou à atualização ou à criação de novos indicadores (MECD, 2014).

Além do mais, essa forma de controle é uma política de Estado, geralmente estabelecida por lei – como se observa no caso do Brasil – e a decisão relacionada aos meios e formas de realização é uma prerrogativa de governo, ou seja:

a avaliação é quase sempre impositiva, sem consulta a professores e muito menos a alunos. A avaliação, por sua vez, é igualmente repetitiva, no sentido de que, ao longo de vários semestres, os alunos fazem avaliações internas e externas, sendo que destas últimas não conhecem os resultados de seus desempenhos e das primeiras têm apenas um escore ou nota sem qualquer tipo de *feedback* que lhes possa servir de orientação (VIANNA 2003, p. 46).

Sendo assim, as avaliações tornam-se exaustivas e geradoras de tensão para alunos e professores, geram polêmicas a favor e contra, mas “não têm maiores consequências na vida dos avaliados” e, desse modo “perdem o seu significado”,

[...] ainda que aos dados, resultantes de comportamentos inteiramente descompromissados, sejam aplicado procedimentos estatísticos complexos que, por sua vez, geram todo um filosofar supostamente baseado em elementos considerados científicos e levam a decisões de repercussão, criando-se, assim, ideias falaciosas em grande parte da sociedade, que, apesar de tudo, passa a acreditar nas conclusões estabelecidas como se verdades absolutas fossem (VIANNA 2003, p. 46).

Para concluir, pode-se afirmar que a generalização global desse modelo de avaliação reduz a importância da avaliação que é feita pelo professor na sala de aula e fortalece o monitoramento externo e centralizado, conforme Freitas, pois:

Nesses sistemas de avaliação, o desempenho individual é subsumido nas estatísticas que lidam, preferencialmente, com tendências globais dos sistemas de ensino ao longo do tempo, *a partir da proficiência média dos alunos*. [...] Há de se reconhecer as falhas nas escolas, mas há de se reconhecer, igualmente, que há falhas nas políticas públicas, no sistema socioeconômico etc. Portanto, esta é uma situação que, à

---

<sup>4</sup>*Trends in International Mathematics and Science Study (TIMSS), Progress in International Reading Literacy Study (PIRLS), Program for International Student Assessment (PISA), Programme for the International Assessment of Adult Competencies (PIAAC) y PISA resolución de problemas.*

espera de soluções mais abrangentes e profundas, só pode ser resolvida por negociação e responsabilização bilateral: escola e sistema. Os governos não podem “posar” de grandes avaliadores, sem olhar para seus pés de barro, para suas políticas, como se não tivessem nada a ver com a realidade educacional do país de ontem e de hoje. (FREITAS, 2007 p. 973-975).

Contudo, não se está afirmando que a avaliação do sistema não seja necessária, pelo contrário, concorda-se com Freitas que esta deve estar articulada com a avaliação institucional e a do professor em sala de aula, pois:

[...] a avaliação de sistema é um instrumento importante para monitoramento das políticas públicas e seus resultados devem ser encaminhados, como subsídio, à escola para que, dentro de um processo de avaliação institucional, ela possa usar estes dados, validá-los e encontrar formas de melhoria. [...] Os três níveis de avaliação (sistema, escola e sala de aula) devem estar trabalhando articuladamente e segundo suas áreas de abrangência (FREITAS, 2007, p. 978-979).

### 1.3. INDICADORES EDUCACIONAIS E ORGANISMOS MULTILATERAIS

Para avaliar os resultados da educação oferecidos pelos organismos multilaterais e governos nacionais, o primeiro passo é a definição de um padrão de qualidade, para que possam ser elaborados os seus indicadores. Neste sentido, a qualidade tornou-se um objeto precioso para os países. Tal como foram determinados indicadores de qualidade para todas as áreas, também foram definidos os da educação. Avaliar alunos, escolas e professores tornou-se uma *peça-chave* para conhecer o desenvolvimento educacional de um país.

Em âmbito nacional e internacional, foram criados muitos programas de pesquisa para avaliar a qualidade da educação, como por exemplo, o *Programme for International Student Assessment* ou Programa Internacional de Avaliação de Estudantes (PISA). Segundo o Instituto Nacional de Pesquisas Internacionais Anísio Teixeira (INEP):

O objetivo do Pisa é produzir indicadores que contribuam para a discussão da qualidade da educação nos países participantes, de modo a subsidiar políticas de melhoria do ensino básico. A avaliação procura verificar até que ponto as escolas de cada país participante estão preparando seus jovens para exercer o papel de cidadãos na sociedade contemporânea (INEP, 2009).

Com a mesma intenção, em 2004, por meio de uma parceria de várias organizações governamentais e não governamentais, foi publicado o documento *Indicadores da Qualidade na Educação*, coordenado pela Ação Educativa<sup>5</sup>, Unicef, PNUD, Inep-MEC. Assim, o objetivo principal desses indicadores de qualidade na educação foi criado “para ajudar a comunidade escolar na avaliação e na qualidade da escola” (AÇÃO EDUCATIVA et al, 2004, p. 5). Tal como consta no documento, indicadores são:

[...] sinais que revelam aspectos de determinada realidade e que podem qualificar algo. Por exemplo, para saber se uma pessoa está doente, usamos vários indicadores: febre, dor, desânimo. Para saber se a economia do país vai bem, utilizamos como indicadores a inflação e a taxa de juros. A variação dos indicadores nos possibilita constatar mudanças (a febre que baixou significa que a pessoa está melhor; a inflação mais baixa no último ano indica que a economia está melhorando). (AÇÃO EDUCATIVA et al, 2004, p.05).

Desse modo, pode-se afirmar que indicador é uma forma de analisar determinado acontecimento durante um tempo determinado, que pode ser medido por horas, dias, semanas, anos, etc. Os indicadores são medidas composta de variáveis, ou seja: “São medições baseadas em mais de um dado e sua construção se dá a partir do somatório de resultados individuais” (CORRADINI, 2012 p. 17).

Todavia, a definição de indicadores relaciona-se com o que se pretende, ou neste caso, com as finalidades da educação e para que serve. No atual contexto, para Corrales, os indicadores revelam que:

A melhoria da qualidade da educação está vinculada a dois objetivos de suma importância para os estados modernos. Em primeiro lugar, a melhoria da qualidade da educação é vista, cada vez mais, como fonte de competitividade econômica internacional para as nações. Em uma economia globalizada os países competem entre si pelos mercados, pelo investimento estrangeiro, pelo desenvolvimento tecnológico, e pela atração de multinacionais. Uma força de trabalho com alto nível de educação é vista como fator de vantagem competitiva nesta concorrência econômica. Em segundo lugar, a educação de alta qualidade passou a ser sinônimo de desenvolvimento nacional autossustentado, não apenas de competitividade internacional (CORRALES, 2000 p. 04).

---

<sup>5</sup>As organizações parceiras nesse estudo são: Unicef, PNUD, Inep, Campanha Nacional pelo Direito à Educação, Cenpec, CNTE, Consed, Fundação Abrinq, Fundescola-MEC, Seif-MEC, Seesp-MEC, Caise-MEC, IBGE, Instituto Pólis, Ipea, Undime e Uncme.

Outro fator importante, para que os indicadores se tornem ferramentas apropriadas e úteis aos objetivos mencionados, é que estes devem ter certas características, como consta nos relatos de Alves:

- a) Objetividade – devem ser preferencialmente quantitativos para reduzir o nível de subjetividade que pode dificultar a mensuração;
- b) Mensurabilidade – devem ser mensuráveis e permitir uma identificação de uma qualidade através de uma escala de valores;
- c) Compreensibilidade – devem informar o desempenho de forma inteligível e traduzir algum significado para seu usuário;
- d) Comparabilidade – devem permitir a comparação da *performance* histórica de uma mesma organização, política ou programa ou entre diferentes organizações, políticas ou programas de diferentes regiões;
- e) Custo – a relação custo/benefício deve ser viável para a realização da avaliação e, consequentemente, para a geração do indicador; (ALVES, 2011, p.45).

Para criar um indicador é necessário estabelecer um conjunto de critérios de avaliação que se relacionam com a concepção de qualidade de um determinado objeto. Souza Júnior esclarece que todo indicador deve ser: *atribuível*, porque ocorrem mudanças nos processos deflagrados pelos atores em evidência; *sensível*, para refletir as mudanças que ocorrem no fenômeno em questão; *confiável*, para obter qualidade no levantamento dos dados; *inteligível*, capaz de mostrar de forma transparente a metodologia de construção do indicador; e *comunicável*, compreensível, para os vários indivíduos e grupos que os utilizam (SOUZA JUNIOR, 200-, p. 9).

Além disso, deve-se averiguar a organização, seja pública ou privada, que realiza a pesquisa: Qual é o status da organização? Quais são os seus objetivos? O que e para que se propõe a medir? Como são definidos os indicadores de qualidade? Quais indicadores devem ser monitorados e com que frequência, para garantir objetividade da pesquisa?

As organizações internacionais, como a OCDE, estabelecem indicadores para medir a qualidade da educação. Para a pesquisa TALIS, aqui tomada como exemplo, foram utilizados 15 indicadores para criar o questionário dos diretores e dos professores. Tal como consta no relatório da OCDE, esses indicadores foram:

- (1) Reconhecimento, *feedback*, recompensa e avaliação de professores;
- (2) Liderança escolar; (3) Práticas de ensino, crenças e atitudes; (4) Qualidade dos professores (experiência, qualificações, responsabilidades); (5) Satisfação e eficácia da educação e da formação;
- (6) Perfil de formação dos professores; (7) Ambiente escolar; (8) Divisão do tempo de trabalho; (9) Frequência e distribuição da educação e da



formação; (10) Satisfação no trabalho e medidas de recursos humanos; (11) Motivação e experiência no início da carreira de novos professores; (12) Perfil dos novos professores; (13) Eficácia dos procedimentos e incentivos no recrutamento e seleção de professores (14) Desgaste dos professores e volume de trabalho; (15) Adequação da oferta de professores e escassez de professores (OCDE, 2009 p. 27).

A partir desses indicadores foram elaborados os questionários aplicados aos professores e aos diretores das escolas selecionadas para a pesquisa (Anexo 1 e 2). O processo de elaboração de questões elaboradas a partir de indicadores é complexo e, portanto, exige um conhecimento especializado da área da Estatística, que a maioria dos profissionais de outras áreas, como também os da educação, desconhece, fazendo uma leitura superficial e, muitas vezes, equivocada dos resultados apresentados em relatórios que, ao final, vão orientar a avaliação e proposição de políticas. Além disso, testes e questionários,

[...] demandam uma complexa cadeia tecnológica que estados e municípios não têm como dominar, o que abre espaço para o crescimento da indústria da avaliação e da tutoria. Se para as corporações interessa o recurso à fixação de “standards” como forma de triar a força de trabalho e monitorar os fluxos de qualificação de mão de obra, além do óbvio controle ideológico da educação, para a indústria educacional os objetivos são mais imediatos e referem-se à conversão da educação em mais espaço mercadológico. Os testes têm seu lugar no mundo educacional como uma ferramenta de pesquisa. O grave problema é que eles foram sequestrados pelo mercado e pelo mundo dos negócios e nele, as suas naturais limitações são ignoradas. [...] os testes associam à sua função de medir, o papel de controle ideológico dos objetivos da educação – mais pelo que excluem do que pelo que incluem – e têm o objetivo de controlar os atores envolvidos no processo educativo. Sem testes, não há responsabilização e meritocracia – teses fundamentais do mercado (FREITAS, 2007 p. 10).

Desse modo, os profissionais que desenvolvem todos os procedimentos técnicos necessários para pesquisas de grande porte são requisitados em empresas privadas que atendem aos interesses dos clientes, ou seja, das agências ou órgãos governamentais que os contratam, que não correspondem obrigatoriamente às necessidades sentidas pelos professores e estudantes. Nesse contexto, a educação passou a ser *presa fácil* de estudos internacionais, que estabelecem seus padrões de qualidade sem resistência, tomando a estatística como principal ferramenta.

### 1.3. PESQUISA TALIS: UM EXEMPLO DE PESQUISA DE LARGA ESCALA

Neste tópico, apresenta-se, como um exemplo de pesquisa de larga escala, a pesquisa *Teaching and Learning International Survey* (TALIS), denominada no Brasil como *Pesquisa Internacional sobre Ensino e Aprendizagem*, estudo realizado pela *Organização de Cooperação e Desenvolvimento Econômico* (OCDE), que avalia as condições de trabalho de professores de países, membros e convidados (OCDE, 2009).

O programa de pesquisa TALIS, desenvolvido em conjunto com membros dos países OCDE e países parceiros, é a primeira pesquisa internacional com foco nas condições de trabalho de professores e no aprendizado das escolas de ensino do nível secundário básico<sup>6</sup> para desenvolver índices relevantes e comparáveis.

Para tanto, é importante conhecer seu contexto, seus objetivos e processo da pesquisa, bem como sua metodologia e os resultados que esta investigação trouxe para contribuir no processo de estudo sobre os professores e suas condições de trabalho, a fim de compreender as intenções e expectativas da OCDE em relação às políticas educacionais da área.

#### 1.3.1 Contexto e objetivos da pesquisa

A pesquisa TALIS, elaborada pela OCDE, faz parte do atual contexto de globalização, no qual os organismos multilaterais ganharam um protagonismo importante na definição de diretrizes e orientação de políticas educacionais dos Estados Nação.

Com a globalização, os países passaram a se preocupar com a definição de padrões de qualidade internacional para a educação, com mais veemência. De tal modo, passam a estabelecer acordos e agendas em comum. Por intermédio de organizações multilaterais, os países passam a se unir para analisar a educação tanto dos países desenvolvidos como dos países em desenvolvimento (DALE, 2002). Nesse sentido, a OCDE elaborou programas de pesquisa internacional, para

---

<sup>6</sup> Nível secundário básico –No Brasil refere-se ao ensino fundamental séries finais (6º ao 9º anos ou 5ª 8ª série).

analisar o processo de ensino e aprendizagem, a partir das necessidades da economia.

A *Organização de Cooperação e Desenvolvimento Econômico* (OCDE) nasceu no período pós-guerra, em meados de 1947, para ajudar os países europeus destruídos pela guerra. Neste contexto, os Estados Unidos da América criou o *Plano Marshall*, com o objetivo de ajudar financeiramente a recuperação dos países aliados europeus (INE, 2006).

Em 16 de abril de 1948, foi criada a *Organização Europeia de Cooperação Econômica* (OECE), com o objetivo de promover a cooperação entre os países membros e coordenar a distribuição dos fundos. Não obstante, a OECE também objetivou “a criação de zonas de mercado livre, a promoção de políticas que melhorassem as condições de trabalho e de produtividade, dentre outras<sup>7</sup>” (INE, 2006).

Os países fundadores da OECE foram a Áustria, Bélgica, Dinamarca, França, Grécia, Islândia, Irlanda, Itália, Luxemburgo, Noruega, Países Baixos, Portugal, Reino Unido, Suécia, Suíça e Turquia, sendo que, em 1955, a Alemanha e, em 1959, a Espanha, também aderiram.

A OECE passou a ser chamada, oficialmente, de *Organização de Cooperação para o Desenvolvimento Econômico* (OCDE) em 03 de setembro 1961, devido à adesão dos Estados Unidos da América e do Canadá, passando a ser não apenas uma organização europeia, mas mundial. Atualmente, a OCDE conta com 34 países membros.

São muitos os objetivos das OCDE, podendo ser resumidos nos seguintes pontos:

- Apoio a um crescimento econômico duradouro, realizando a maior expansão possível da economia, do emprego e da qualidade de vida dos países membros, mantendo a estabilidade financeira e contribuindo consequentemente para o desenvolvimento equilibrado da economia mundial;
- Favorecer a expansão do comércio internacional apoiado numa base não discriminatória;
- Combater a corrupção no campo da política mundial;
- Ajudar governos e sociedade no processo de globalização;
- Ajudar no desenvolvimento além-fronteiras, estabelecendo parcerias com o setor empresarial e a sociedade civil, através da *Business Industry Advisory Committee to the OECD* (BIAC), da *Trade Union Advisory Committee to the OECD* (TUAC) e de organizações não governamentais;

---

<sup>7</sup>Tradução livre.

- Constituir uma fonte de dados estatísticos, cuja base cobre áreas que vão desde indicadores econômicos até os relacionados com educação, emprego, energia, comércio e saúde (INE 2006, p. 25).

Importante destacar que todos os países membros contribuem financeiramente para a manutenção da organização, de acordo com o tamanho da economia de cada um. Assim, os Estados Unidos da América cooperam com 25% do orçamento total, constituindo-se no maior contribuinte. Ressalta-se, também, que existem 250 comissões e grupos de trabalho, que reúnem representantes dos países membros para debater e avaliar os progressos feitos em domínio políticos específicos (economia, saúde educação, etc.) e definir padrões e indicadores de qualidade (INE 2006).

Assim sendo, a educação passou a ser alvo de estudos internacionais para estabelecer seus padrões de qualidade. A preocupação com a questão docente foi abordada em 2005 pelo estudo *Atrair, Formar e Manter Professores Eficazes* (OCDE, 2005, p.01). Esse trabalho, que iniciou em abril de 2002 e foi publicado em junho de 2005, compreende um estudo analítico em um temático. Do estudo analítico, participaram 25 países: Alemanha, Austrália, Áustria, Bélgica (Comunidade Flamengo), Bélgica (Comunidade francesa), Canadá (Quebec), Chile, Coréia, Dinamarca, Espanha, Estados Unidos, Finlândia, França, Grécia, Hungria, Irlanda, Israel, Itália, Japão, México, Noruega, Holanda, República Eslovaca, Reino Unido, Suécia e Suíça. E do estudo temático: Alemanha, Áustria, Bélgica (Comunidade Francesa) Coréia, Espanha, Hungria, Itália, Suécia e Suíça. Tinha como intenção, auxiliar os governos a idealizar e aperfeiçoar as políticas relacionadas com os professores e melhorar o ensino e a aprendizagem, com o objetivo de:

Fornecer uma análise abrangente internacional sobre: os progressos e tendências na profissão docente; os dados sobre os fatores-chave para atrair, formar e reter professores de qualidade; a indicação políticas e práticas inovadoras e eficazes; as opções para formulação de políticas a respeito dos professores; e a definição de prioridades para futuros trabalhos (TROJAN, 2009, p.02).

O documento publicado sobre os estudos *Ensinar importa: Atrair, Formar e Manter Professores Eficazes* levantou uma serie de problemas relacionados com a docência, entre os quais o envelhecimento da profissão, a inadequação da formação inicial e continuada e das políticas decorrentes das reformas de Estado, realizadas nos anos 1990 (OCDE, 2005).

Desse relatório, a OCDE chegou à conclusão de que era necessário obter melhores informações, de nível nacional e internacional, sobre os professores. Com esse estudo de revisão e falhas encontradas nos índices e prioridades, é que se criaram os instrumentos utilizados no desenvolvimento da pesquisa TALIS. De tal modo, a pesquisa foi desenvolvida para analisar questões, tais como as práticas de ensino que estão sendo aplicadas, as crenças e atitudes dos professores, o clima escolar, e como estes fatores influem na qualidade do processo de ensino e aprendizagem. Ou seja, o programa surgiu para estudar, sobretudo, as condições de formação e de trabalho dos docentes e de suas escolas (OCDE, 2009).

Portanto, a pesquisa tem como objetivo contribuir para a melhoria dos padrões educacionais e “conhecer e analisar aspectos que possam contribuir para a qualidade de ensino” (CORADINI, 2012, p.92), criar ou contribuir na melhoria das políticas educacionais, e também validar as estatísticas – nesse contexto, as de larga escala.

Para entender melhor o programa TALIS, é importante conhecer melhor sua raiz, sua ideia de analisar as condições de ensino e aprendizagem oferecidas pelos sistemas educacionais. TALIS foi desenvolvido como parte do projeto dos *Indicators of Education Systems* (INES) (Indicadores de Sistemas de Educação) da OECD. Esses indicadores foram coletados durante 20 anos, formando a base de conhecimentos para analisar o desempenho dos sistemas educacionais dos países vinculados à OCDE (2008, p. 12).

A estrutura conceitual original e a criação do programa TALIS surgiu através de uma força tarefa que incluiu peritos do *Network A* (resultados de aprendizagem) do INES e da *Network C* (ambiente de aprendizagem e organização escolar). Essa ação teve como objetivos desenvolver e criar um banco de dados direcionados a professores e diretores de escolas dos países membros e convidados da OCDE, para analisar, indicar falhas e aperfeiçoar os indicadores INES (OCDE, 2009, p.21).

O principal produto do INES é o evento anual *Uma mirada da Educação*. Em 2003 na cidade de Dubai, os ministros da educação levantaram aspectos importantes em relação à necessidade de qualidade e de como o ensino influencia na aprendizagem (OCDE, 2009). Portanto, o objetivo geral das pesquisas TALIS é:

[...] promover, de forma efetiva, tanto em questão de tempo quanto de custo, indicadores internacionais fortes e uma análise relevante das políticas sobre professores e ensino para ajudar os países a revisar e

desenvolver políticas que criam as condições para um ensino eficiente. Análises entre países proporcionam a oportunidade de comparar nações que passam por desafios similares e também aprender diferentes abordagens políticas e seus impactos no ambiente de aprendizagem das escolas (OCDE, 2009 p. 13).

Outro aspecto relevante para a implantação do programa foi o engajamento dos países membros e parceiros da OCDE, que criaram uma comissão para representar os países e estabelecer os objetivos da pesquisa e constituir os padrões para o recolhimento dos dados. A partir das conclusões dessas reuniões, foram contratadas empresas especializadas para realização de cada etapa da pesquisa. Assim sendo: (1) A *Trades Union Advisory Council* (TUAC) foi importante para o desenvolvimento e implementação do programa; (2) O grupo *Instrument Development Expert Group* (IDEG) foi estabelecido para traduzir as prioridades das políticas selecionadas pelos países, em forma de questionário; (3) a *National Project Managers* (NPMs) e a *National Data Managers* (NDMs), tiveram um papel significativo para implementar o programa em nível nacional, onde a NPMs ajudou a validar os questionários, assegurar a cooperação das escolas, gerenciar o recolhimento nacional dos dados, processar e verificar os resultados da TALIS e a NDMs coordenou o processamentos e formação dos dados; (4) o *Data Processing Centre of International Association for the Evaluation of Educational Achievement* (IEA) gerenciou a implantação em nível internacional, verificou a tradução dos questionários e exerceu o controle de qualidade; (5) a *Statistics Canadá*, sub contratada da IEA, desenvolveu a estatística aplicada no programa (OCDE, 2009).

Após a realização da pesquisa, foi elaborado o primeiro relatório, que foi publicado em 2009<sup>8</sup>. O relatório TALIS está subdividido em sete capítulos: o primeiro capítulo apresenta uma visão ampla do programa; o segundo apresenta as características dos professores e das escolas onde trabalham; o terceiro apresenta o desenvolvimento profissional dos professores; o quarto analisa as práticas, crenças e atitudes dos professores; o quinto, a avaliação e o *feedback* que os professores recebem; o sexto opera sobre a liderança escolar; o sétimo, sobre clima disciplinar e autoeficácia dos professores, além da modelagem estatística (OCDE, 2010 p. 17-18). Para analisar os resultados obtidos e o uso da estatística aplicada a essa investigação é necessário conhecer a metodologia empregada em todo o processo.

---

<sup>8</sup>A segunda rodada da pesquisa foi realizada em 2012 e os relatórios publicados em 2014. Estão disponíveis no site do INEP: <http://portal.inep.gov.br/web/talis>.

### 1.3.2. Metodologia utilizada na pesquisa TALIS

O foco da pesquisa TALIS é o ensino de nível secundário básico – que, no Brasil, refere-se às séries finais do ensino fundamental –, seus professores e diretores das escolas, com o intuito de, segundo a OCDE, promover índices relevantes e análises dos seguintes aspectos-chave do ensino: o papel e a função da liderança escolar; como o trabalho dos professores é apreciado e o *feedback* que eles recebem; desenvolvimento profissional dos professores; e crenças dos professores e atitudes em relação ao ensino e suas práticas pedagógicas (OCDE, 2009).

TALIS foi pensado como um programa de pesquisas. O relatório publicado em 2009 apresenta os resultados iniciais da primeira rodada do TALIS, a qual foi implementada em 2007-08. Ou seja, foi concebido como uma sequência de pesquisas, as quais, com o tempo, deverão investigar professores de todas as etapas do ensino<sup>9</sup>.

Por decisão dos países participantes, o foco da primeira rodada da TALIS referiu-se aos professores e diretores da educação secundária básica (nível 2 ISCED<sup>10</sup> 1997). Foram preparados questionários separados para professores e diretores das escolas, para explorar as questões analíticas e sobre políticas sugeridas pelos países participantes. Foram realizados estudos para alcançar validade cultural e linguística dos instrumentos de pesquisa, aplicados tanto para as traduções quanto para as amostras de dados coletados (OCDE 2009).

A população-alvo é constituída de professores e diretores do ensino secundário básico, de escolas públicas e privadas dos países participantes. O tamanho mínimo da amostra foi definido em 200 escolas por país e 20 professores em cada escola, questão que será abordada com mais detalhe, mais adiante. No Brasil, a amostra estudada foi de 400 escolas correspondendo a 7.161 professores;

---

<sup>9</sup> Entretanto, a segunda rodada manteve a abordagem no ensino secundário básico (OCDE, 2014).

<sup>10</sup> ISCED – *International Standard Classification of Education*: Classificação Internacional de Educação de 1997. É um instrumento internacional para identificar a equivalência ou similaridade dos níveis de educação entre os países. Estabelece seis níveis de educação: Educação pré-primária (ISCED nível 0); Educação primária (ISCED nível 1); Educação secundária (básica) (ISCED nível 2); Educação secundária (superior) (ISCED nível 3); Educação pós-secundária de nível não terciário (ISCED nível 4); e a Educação de nível 5, que representa a educação terciária e está dividido entre os níveis 5A e 5B. O nível 5A refere-se à graduação, Bacharelado ou Licenciatura; o nível 5B, geralmente, refere-se a programas mais curtos; e o nível 6 representa educação de nível terciário com pesquisa avançada, como por exemplo, doutorado ou PhD (OCDE 2009).

a estratificação explícita foi organizada por tipo de escola (privada, municipal ou estadual); e a estratificação implícita foi organizada por região (Norte, Nordeste, Centro Oeste, Sul e Sudeste), localização (urbana e rural) e estado federal (27 estados).

No México, a amostra estudada foi de 200 escolas, com 4.164 professores; e a estratificação explícita foi organizada por tipo de escolas (geral particular e técnica); e a estratificação implícita foi organizada nos 32 estados mexicanos.

Na Espanha, o sistema de ensino está sob a responsabilidade de 18 comunidades autônomas. A população alvo da pesquisa abrange 16 comunidades, porque as de *La Rioja* e *Ilhas Canárias* não participaram. A amostra desse país foi de 200 escolas com 4000 professores no total, em conformidade com a regra estabelecida. A estratificação explícita, ou seja, a definição das escolas participantes foi de responsabilidade das comunidades autônomas envolvidas; e a estratificação implícita foi definida por tipo de escola (públicas e privadas) (OCDE 2009).

Para realizar estudos comparados entre os países citados ou estudos intranacionais a partir desta pesquisa, deve-se fazer um levantamento mais detalhado sobre as possibilidades dessas amostras. Por exemplo, pode-se, a priori, verificar que o número de escolas e professores investigados é pequeno em relação ao universo total, inviabilizando comparações intranacionais.

Ainda sobre a questão da amostra, é importante destacar que nem tudo que é selecionado deve ser coletado. Como escolas com até 3 professores e escolas localizadas em regiões geograficamente remotas incidem em um trabalho difícil, caro e estatisticamente ineficiente, foi determinado que alguns países pudessem excluí-las da amostra, sem danificar a qualidade da pesquisa e criar uma nova população de investigação, sem prejudicar a análise final. Nesses casos, as razões para exclusão foram todas documentadas (OCDE, 2009).

As datas de realização da coleta de dados foram outubro e novembro de 2007, para os países do hemisfério sul; e março a maio de 2008, para os países do hemisfério norte, tendo em vista as diferenças de calendário escolar. Os respondentes tiveram 45 minutos para completar os questionários e foram preenchidos em papel ou *online* (OCDE, 2009).

Portanto, somente foram considerados para o estudo, os países que obtiveram respostas de 75% das escolas da amostra e as escolas com 50% dos



professores dos questionários respondidos pelos professores, com o intuito de atingir 75% da resposta de todos os professores da amostra de cada país. Por não atender aos padrões estabelecidos, não foram considerados os dados da Holanda para efeitos de análise geral.

Entretanto, o desenvolvimento de uma pesquisa, para determinar os requisitos necessários para idealizar os padrões de resultado desejados, vai além das fronteiras do senso comum. Na educação, por exemplo, não se pode admitir a ideia de que somente os professores determinam o padrão de qualidade no ensino, pois o processo educativo é complexo e sujeito a inúmeros fatores, relacionados com o contexto do país, da região e do próprio aluno. Portanto, faz-se necessário correlacionar a variável “professor” com outros fatores que interferem nesse universo os quais devem ser investigados para garantir a validade da pesquisa e contribuir para melhor compreender a educação. No que concerne às condições para o trabalho docente:

Há outros elementos igualmente importantes – como a valorização social da profissão, os salários, as condições de trabalho, a infraestrutura das escolas, as formas de organização do trabalho escolar, a carreira – que devem fazer parte de uma política geral de apoio aos docentes. São múltiplos os fatores que não podem ser esquecidos, nem desconsiderados no delineamento de políticas para os professores (GATTI, et al. 2011 p. 15).

Para conhecer e compreender a metodologia de uma pesquisa faz-se necessário identificar a essência organizadora do processo, que decorre do princípio teórico escolhido. Tal princípio:

[...] opera como a expressão, no pensamento, do mais fundamental, do mais real que existe e que, portanto, é o polo oposto da especulação. É a expressão articulada do conjunto de determinações – fundamentos – de uma realidade historicamente dada, das circunstâncias recorrentes e mutantes em suas recorrências, nas quais consiste uma determinada história (ROCHABRÚN, 1974 p. 09)<sup>11</sup>.

Na TALIS, o princípio teórico adotado para examinar *as crenças, as atitudes e práticas* dos professores e comparar países e escolas, fundamenta-se no construtivismo. Nesse sentido, a pesquisa utiliza dois conceitos de prática pedagógica – construtivista e transmissão direta – para classificar os fatores

---

<sup>11</sup> Tradução da Profª Drª Maria Regina Michelotto.

relacionados às crenças e entendimentos básicos dos professores sobre a natureza do ensino e da aprendizagem (OCDE, 2009 p. 89-90).

Na pesquisa entende-se como perspectiva construtivista aquela que “foca no estudante como um participante ativo no processo de aquisição do conhecimento [...]; [dá] aos estudantes a chance de desenvolver soluções aos problemas [...] e [permite] que os alunos tenham um papel ativo nas atividades”. Em contraste, a perspectiva da transmissão direta do ensino e aprendizagem do estudante, sugere que o papel do professor é “comunicar o conhecimento de modo claro e estruturado [...]; fornecer aos estudantes problemas claros e ‘resolvíveis’; [...] e garantir a calma e a concentração na sala de aula” (OCDE, 2009 p.93). A partir dessa explicação, pode-se concluir que a denominada prática de transmissão direta é aquela que conhecemos como tradicional – pautando-se na transmissão de conhecimentos e na centralidade da ação do professor – e, por outro lado, a construtivista – pautada na atividade e na experimentação, cuja centralidade está na ação do aluno (SAVIANI, 2005).

Para analisar as perguntas elaboradas e os resultados apresentados pelo informe TALIS, deve-se levar em consideração esse ponto de vista. Neste sentido, os indicadores foram elaborados reforçando as práticas construtivistas como as mais adequadas. Por exemplo, ao responder às questões citadas, o professor tende a concordar com as práticas que dão ao aluno “a chance de desenvolver soluções aos problemas” e não àquelas que enfatizam o papel do professor na apresentação das respostas certas.

No que se refere ao estudo das características dos professores e das escolas onde trabalham e sobre o desenvolvimento profissional dos professores, a pesquisa TALIS utiliza a média e o desvio padrão, que se refere às medidas de tendência central e de dispersão. Nesse caso, são tabuladas as características apresentadas pelos respondentes – por exemplo: idade, tempo de experiência, número de alunos por sala, tipo de atividade de desenvolvimento profissional, etc. – que são registradas em valores percentuais (SPIEGEL, 1977). Entretanto, apesar do desvio padrão – ou seja, a variação que existe em torno da média – ser apresentado nas tabelas, o documento não os considera para o diagnóstico das questões. Esta é mais uma falha no processo de análise dos dados, que será estudada com mais profundidade mais adiante.

### 1.3.3 Questionários da Pesquisa TALIS

Um questionário possui a importância de buscar dados sobre determinado estudo e levantar problemas sobre os acontecimentos em torno do que se está estudando, bem como permite recolher amostras dos conhecimentos, atitudes, valores e comportamentos sobre o objeto da pesquisa. E para compreender melhor, Amaro et. al (2005) explicam que existem três importantes tipos de questionários, desenvolvidos para atender diferentes tipos de questões: resposta aberta, resposta fechada e misto. Como mostra o Quadro1 com as vantagens e desvantagens para cada determinado tipo de questão.

Quadro1: VANTAGENS E DESVANTAGENS DAS RESPOSTAS ABERTAS E FECHADAS

| TIPO DE QUESTÃO           | VANTAGENS  | DESVANTAGENS   |
|---------------------------|--|--|
| <b>Respostas abertas</b>  | <ul style="list-style-type: none"> <li>▪ Preza o pensamento livre e a originalidade;</li> <li>▪ Surgem respostas mais variadas;</li> <li>▪ Respostas mais representativas e fiéis à opinião do inquirido;</li> <li>▪ O inquirido concentra-se mais sobre a questão;</li> <li>▪ Vantajoso para o investigador, pois permite recolher variada informação sobre o tema em questão.</li> </ul> | <ul style="list-style-type: none"> <li>▪ Dificuldade para organizar e categorizar as respostas;</li> <li>▪ Requer mais tempo para responder às questões;</li> <li>▪ Muitas vezes, a caligrafia é ilegível;</li> <li>▪ Em caso de baixo nível de instrução dos inquiridos, as respostas podem não representar a opinião real do próprio.</li> </ul>   |
| <b>Respostas fechadas</b> | <ul style="list-style-type: none"> <li>▪ Rapidez e facilidade de resposta;</li> <li>▪ Maior uniformidade, rapidez e simplificação na análise das respostas;</li> <li>▪ Facilita a categorização das respostas para posterior análise;</li> <li>▪ Permite contextualizar melhor a questão.</li> </ul>   | <ul style="list-style-type: none"> <li>▪ Dificuldade do pesquisador para elaborar as respostas possíveis a uma determinada questão;</li> <li>▪ Não estimula a originalidade e a variedade de resposta;</li> <li>▪ Não preza uma elevada concentração do inquirido sobre o assunto em questão;</li> <li>▪ O inquirido pode optar por uma resposta que se aproxima mais da sua opinião, não sendo esta uma representação fiel da realidade.</li> </ul> |

Fonte: Adaptado de A arte de fazer questionários, AMARO, et. al., 2005

As questões de resposta aberta não apresentam alternativas e são conhecidas na educação como questões qualitativas, que permitem ao pesquisado expressar suas ideias de uma forma livre, construindo as respostas com suas próprias palavras. Aquelas com respostas fechadas, na educação, são conhecidas como questões quantitativas, que admitem selecionar apenas a opção mais adequada à sua opinião. E o questionário misto, que é composto de questões de respostas abertas e fechadas (AMARO et. al, 2005).

Para a pesquisa TALIS, foram elaborados dois questionários<sup>12</sup>, com questões de respostas fechadas: um para diretores e outro para os professores das escolas investigadas. O questionário dos diretores contou com 37 perguntas, dividido em 4 sessões, com o objetivo de levantar dados sobre: *informações básicas* sobre o diretor, sua escolaridade e sua atuação na direção da escola, abarcando 7 perguntas; *informações básicas sobre a escola*, englobando o tipo de escola, recursos, perfil da comunidade a qual pertence, número de funcionários, número de matrículas e admissão dos alunos na escola, com um total de 7 perguntas; *gestão escolar*, com perguntas sobre o tipo de gestão da escola, importância de seu trabalho na escola, atividades administrativas, auto avaliação da escola, aspectos das avaliações internas e externas realizadas e sua influência no processo de ensino e aprendizagem, com 8 questões; *avaliação do professor*, sobre a apreciação do trabalho do professor pelo diretor de modo formal e objetivo (como parte de um sistema de gestão de desempenho), até uma abordagem mais informal e subjetiva (por exemplo, mediante conversas informais com os professores), envolvendo 6 questões; *recursos da escola*, que faz levantamento de dados sobre a qualidade do ensino (por exemplo, se os alunos estão sendo prejudicados, por algum motivo), autonomia da escola e sobre o ensino e a aprendizagem, envolvendo 9 questões (OCDE 2009).

---

<sup>12</sup> Os questionários do diretor, em português, estão disponíveis para consulta no site do INEP. Da primeira rodada em:  
[http://download.inep.gov.br/download/internacional/talis/talis\\_questionario\\_do\\_diretor\\_br.pdf](http://download.inep.gov.br/download/internacional/talis/talis_questionario_do_diretor_br.pdf)  
 E da segunda rodada em:  
[http://download.inep.gov.br/acoes\\_internacionais/pesquisa\\_talis/2013/diretor\\_quest\\_frequencias.pdf](http://download.inep.gov.br/acoes_internacionais/pesquisa_talis/2013/diretor_quest_frequencias.pdf)  
 Os questionários do professor, em:  
[http://download.inep.gov.br/download/internacional/talis/talis\\_questionario\\_do\\_professor\\_br.pdf](http://download.inep.gov.br/download/internacional/talis/talis_questionario_do_professor_br.pdf)  
 E da segunda rodada em:  
[http://download.inep.gov.br/acoes\\_internacionais/pesquisa\\_talis/2013/professor\\_quest\\_frequencias.p](http://download.inep.gov.br/acoes_internacionais/pesquisa_talis/2013/professor_quest_frequencias.pdf)  
 df

O questionário dos professores está dividido em 5 sessões, abrangendo 43 perguntas, com respostas fechadas, com objetivo de levantar dados sobre: *informações básicas*, com perguntas referentes ao professor, sua escolaridade e o seu tempo de docência, contendo nesse tópico 10 questões; *desenvolvimento profissional*, que é definido como as atividades que desenvolvem habilidades, conhecimento, expertise e outras características do professor, com 10 questões; *avaliação do professor e feedback*, na qual está contida a avaliação do professor sobre o seu trabalho e sobre a gestão da escola e atuação do diretor, bem como a avaliação realizada sobre seu trabalho como docente, realizada por meio de abordagem formal (como parte de um sistema de gestão com critérios pré-estabelecidos), ou até uma abordagem informal (por intermédio de um relatório escrito, ou informalmente, como por exemplo, em discussões com o professor) e o Feedback, como comunicação dos resultados da apreciação do trabalho do professor, contando com 8 questões; *práticas de ensino, crenças e atitudes*, desenvolvido para levantar dados referente à metodologia de ensino e aprendizagem, como é o professor em sala de aula, como é a gestão da escola em relação ao professores, e qual o tipo de disciplina lecionada na referida escola, contendo 5 questões; e, finalmente, *o seu ensino para uma turma específica nesta escola*, onde se faz levantamento sobre a turma, o tempo de formação na matéria lecionada, bem como o número de aluno por turma, e como são as habilidades e comportamentos dos alunos nessa matéria específica, totalizando 10 questões (OCDE, 2009).

Cabe ressaltar que, discorrer sobre a pesquisa TALIS, sua finalidade, definição de sua amostra, a opção de questionário fechado e seu conteúdo se justificam porque esta servirá de referência para analisar a aplicação da Estatística e suas possibilidades para avaliar políticas educacionais. É importante destacar que as análises presentes no relatório (OCDE, 2009) limitam-se à estatística descritiva e não foi feita análise inferencial, que é fundamental para interpretação dos dados. Em geral, os profissionais do campo educacional desconhecem esta diferença, os métodos e as normas da estatística. Por isso, é fundamental conhecer Estatística para selecionar o método mais adequado e não parar na mera descrição. Assim sendo, a seguir será feito um recorrido histórico sobre a Estatística e o seu uso na educação.

## 2. FUNDAMENTOS TEÓRICOS DA ESTATÍSTICA

"É notável que tal ciência, que começou nos estudos sobre jogos de azar, tenha alcançado os mais altos níveis do conhecimento humano" (Laplace).

Antes de explicar a estatística como ferramenta de pesquisa, cabe explicar o que se entende pelo termo *pesquisa*. A pesquisa se relaciona com busca, procura ou investigação, ou seja, com a elaboração de novos conhecimentos. Segundo Luna (2007, p. 15): "Essencialmente, pesquisa visa à produção de conhecimento novo, relevante, teórica e socialmente fidedigno". Ou seja, os resultados de uma pesquisa apresentam interpretações sobre o problema investigado, cujos fundamentos teóricos devem apresentar evidências passíveis de aceitação por parte da sociedade na qual a realidade em foco está inserida.

Assim, para estudar a estatística devem-se buscar seus pressupostos, fundamentando-os nos principais especialistas da área. Nesse trabalho, selecionamos Siegel (1975), Medeiros (2007), Reis (2008), Sganzerla (2001) Bayer et al., (2004), Bolfarine e Bussab (2005), Memória (2004) entre outros, para elucidar essa importante ferramenta, que através de suas análises pode gerar mudanças de padrões na sociedade.

Em primeiro lugar, é necessário estabelecer um marco teórico dessa ferramenta, a fim de elucidar os conceitos fundamentais que serão utilizados no presente estudo. Desse modo, partimos do estudo da história, dos conceitos, princípios e métodos da Estatística para, em seguida, analisar como pode ser utilizada no âmbito da educação, finalizando com indicação das suas contribuições e problemas.

### 2.1 A ESTATÍSTICA NA HISTORIA

As estatísticas estão presentes em inúmeros meios de comunicação, como revistas, jornais, etc., para medir e explicar a vida cotidiana das pessoas. No atual mundo globalizado e informatizado, esta ferramenta cresceu e ocupou um espaço que outrora não se imaginava. É tratada como *vilão* por alguns e *mocinho* por outros. O problema está em quem e como se interpreta e se comprova uma hipótese. Uma boa interpretação, aliada ao bom senso, é que torna uma estatística eficaz, independentemente do resultado comprovar ou refutar a hipótese inicial.

Senra explica, de forma objetiva, o significado da estatística de acordo com a função que ocupa para explicar os fenômenos:

As estatísticas são expressões numéricas de coletivos idealizados cientificamente (por algum princípio de equivalência); por usarem números, atendem a tradição científica ocidental, objetiva e universal. As estatísticas configuram como individualidades individualizadas (individualizações) as individualidades, que, por serem variadas e múltiplas, são, naturalmente, ingovernáveis e incontroláveis. Pelas estatísticas, as individualidades individualizadas tornam-se governáveis e controláveis (SENRA, 2009 p.18).

Mas, para entender melhor esse campo, é fundamental conhecer sua origem, desenvolvimento e como foi se articulando com a vida humana ao longo da história. Autores como Medeiros (2007), Bayer et al., (2004), Sganzerla (2001) e Memoria (2004) fazem um resgate de como a Estatística era usada, desde os povos da antiguidade, passando pela era cristã e acompanham o seu percurso no século XVIII, quando finalmente, o alemão Gottfried Achemmel a denominou como ciência. E, sobretudo nos escritos de Bayer et al., (2004), encontram-se referências históricas do tema no contexto brasileiro.

Segundo Sganzerla (2001, p.11), inicialmente a Estatística estava ligada apenas à descrição de fatos e que:

A representação gráfica 'quantitativa' tem sua origem na construção de cartas geográficas, que datam em torno de 6000 anos, gravadas em argila na Mesopotâmia. [...] Medidas de valor central podem ser encontradas já nos séculos V e VI antes de Cristo. Wallis e Roberts citam um exemplo do uso da moda ou valor mais popular, no século V antes de Cristo. A prática de formar algum tipo de valor central é muito antiga; Walker descobriu indícios do uso da média já no tempo de Pitágoras.

No que concerne ao uso da estatística pelos povos da antiguidade, Medeiros (2007) situa: (1) no ano de 3050 a.C., o primeiro levantamento estatístico de que se tem conhecimento se deve a Heródoto e se refere a um estudo da riqueza da população do Egito, cuja finalidade era averiguar quais eram os recursos humanos e econômicos disponíveis para a construção das pirâmides; (2) no ano de 2238 a.C., o imperador chinês Yao ordenou a realização de uma estatística com fins industriais e comerciais; (3) no ano de 1400 a.C., o famoso faraó egípcio Ramsés II ordenou um levantamento das terras do Egito (MEMORIA, 2004).

Na Grécia, também aparecem registros de levantamentos estatísticos. Mirshawka relata que: “Em diálogo platônico, Sócrates diz a Glauco a necessidade que é para os homens do governo o desenvolvimento das indagações estatísticas” (MIRSHAWKA, 1975, p. 8).

Sobre a prática do recenseamento, um dos acontecimentos de maior relevância histórica para a Estatística ocorreu na Era Cristã, conforme é narrado na Bíblia:

Naquele tempo o imperador César Augusto mandou uma ordem para todos os povos do Império. Todas as pessoas deviam se registrar a fim de ser feita uma contagem da população. Quando foi feito esse primeiro recenseamento, Cirênio era governador da Síria. Então todos foram se registrar, cada um na sua própria cidade. Por isso José foi de Nazaré, na Galileia, para a região da Judeia, a uma cidade chamada Belém, onde tinha nascido o rei Davi. José foi registrar-se lá porque era descendente de Davi. Levou consigo Maria, com quem tinha casamento contratado. Ela estava grávida, e aconteceu que, enquanto se achavam em Belém, chegou o tempo de a criança nascer. Então Maria deu à luz ao seu primeiro filho. Enrolou o menino em panos e o deitou numa manjedoura, pois não havia lugar para eles na pensão (BÍBLIA, N.T. Lucas, 2:1-7).

Por ventura, um acontecimento fortuito, que revela de relevância significativa da estatística para o controle e registro do censo populacional da época. Fato importante que mostra a trajetória dessa ciência e demarca o início da era cristã, definindo a contagem dos anos e séculos, com a denominação de antes e depois de Cristo. Foi depois do início da era cristã que se encontram estudos estatísticos descritivos mais elaborados. Assim,

Os gráficos estatísticos são [...] utilizados por Leonardo da Vinci para analisar a rapidez dos objetos em movimento, aprofundado por René Descartes em 1637, num apêndice do *Discours de La Méthode* (Discurso do Método), na sua obra intitulada *La Géométrie* (A Geometria) (SGANZERLA, 2001, p.11).

Bayer et al (2004), Memória (2004) e Sganzerla (2001) narram outros fatos marcantes da era d.C.: (1) no ano de 620, em Constantinopla, surgiu o primeiro *bureau* de estatística; (2) no ano de 1654, os princípios do cálculo de probabilidades para solucionar problemas relacionados a jogos de azar, no salões da França, foi estabelecido por Blaise Pascal e Pierre de Fermat, considerados pilares da Estatística; (3) em 1708, foi criado o primeiro curso de estatística na Universidade de IENA, Alemanha; (4) no século XVIII, Pierre Simon, Marquês de Laplace, desenvolveu a *Teoria Central do Limite*, segundo a qual “qualquer soma ou média



de variáveis aleatórias tem, para um grande número de termos, uma distribuição aproximadamente normal” (MEMÓRIA, 2004 p. 17); (5) nesse mesmo século, Carl Friedrich Gauss desenvolveu a *curva de erros*, conhecida como *normal padrão*; (6) Galton introduziu os termos mediana, embora realmente foi Gauss quem já havia sugerido esse termo, e o percentil fora desenvolvido para finalidades intimamente relacionadas ao seu uso moderno, em 1885.(7) a variância nasceu, com os mínimos quadrados, e em 1893, Karl Pearson desenvolveu o coeficiente de correlação e introduziu o termo *desvio padrão*, representado por  $\sigma$ .; (8) em 1918, Ronald Aylmer Fischer desenvolveu, entre outros métodos, o *método de máxima verossimilhança* (MEMÓRIA, 2004). Tais formulações são utilizadas para os estudos estatísticos até os dias de hoje.

O mais apropriado significado da estatística teve origem nas suas aplicações, pois essa ciência interage com todas as outras, devido a sua natureza que se relaciona ao uso de dados. Daí a sua importância como instrumento fundamental para auxiliar na pesquisa científica (MEMÓRIA, 2004).

Os levantamentos por amostragem desenvolveram-se após as contribuições da teoria estatística como o principal método estatístico utilizado nas ciências econômicas e sociais. Kier no final do século XIX apresentou a ideia de amostra representativa, como uma representação da população em miniatura, e propôs o método que é conhecido hoje como o método de estratificação, levando em conta fatores geográficos, sociais e econômicos. O período entre 1895 e 1925 marcou a mudança nas estatísticas governamentais de plena confiança na cobertura total fornecida pelos censos para obtenção de informações adicionais nos inquéritos parciais. Este período demarca o que é conhecido hoje como Estatística Moderna (MEMÓRIA, 2004).

Posteriormente, foram ampliados esses estudos em diversos países. Nos Estados Unidos, tais atividades foram exercidas pelo *Bureau of the Census* (1790), com trabalho pioneiro no desenvolvimento, construção e aplicação do equipamento de processamento de dados em cartões perfurados, conhecidos como cartões *Hollerith*<sup>13</sup>. Atualmente, esse processamento é feito por computação eletrônica (MEMÓRIA, 2004).

---

<sup>13</sup>Cartões *Hollerith*:Um tabulador estatístico construído pelo norte americano Herman Hollerith, para acelerar o processamento das estatísticas para o censo dos Estados Unidos de 1890. Esses cartões eram do tamanho das notas de 1 dólar. Cada cartão tinha doze fileiras de vinte orifícios, que eram

Em Bayer et al., (2004), também são retratados fatos importantes no que se refere ao uso da estatística no Brasil, quais sejam: (1) no ano de 1872, a realização do primeiro censo geral da população brasileira, feito por José Maria da Silva Paranhos, conhecido como Visconde do Rio Branco; (2) em 1936, a criação do Instituto Brasileiro de Geografia e Estatística (IBGE); (3) em 1953, o início do ensino de Estatística no Brasil, em duas escolas: a *Escola Nacional de Ciências Estatística* (ENCE) no Rio de Janeiro e a *Escola de Estatística* da Bahia.

## 2.2. EDUCAÇÃO E ESTATÍSTICA

O estudo sistemático da Estatística apareceu nos anos de 1660 na Alemanha, na disciplina Ciência do Estado, para entender o seu sistema e organização. Nos anos de 1777, com o mesmo objetivo, foi introduzido nas universidades da Áustria e na Itália, em Pávia (1814) e Pádua (1815), fazendo parte dos cursos de *Ciências de Leis e Políticas*. A partir de 1849, nas universidades belgas, passou a integrar os currículos de Filosofia e Letras, como ensino da *Aritmética Social*. Esse novo enfoque dado à disciplina foi devido à relevância que a estatística obteve como ferramenta para descrever a situação econômica dos países europeus, após a revolução francesa. Nos Estados Unidos, foi introduzida, em 1845, como instrumento para estudos sobre aspectos morais e intelectuais do homem. (LOPES, 1998).

No final do século XIX e início do século XX, a estatística não mais se limita apenas ao estudo da Demografia e da Economia, e sim, passa a transitar em todas as outras áreas como uma ferramenta importante para a análise de dados em Biologia, Medicina, Física, Psicologia, Indústria, Comércio, Meteorologia, Educação etc., a partir das descobertas de Ronald Fischer e Karl Person (MEDEIROS, 2007).

Nos Estados Unidos, o *International Statistical Institute* (ISI), criado em 1885, e a *United Nations Educational, Scientific and Cultural Organization* (UNESCO), como órgão da *Organização das Nações Unidas* (ONU), criaram, em 1949, um comitê de educação dentro do ISI, com objetivo de promover a educação estatística

---

perfurados para registrar dados sobre idade, sexo, país natal, número de filhos, profissão, estado civil e tudo o mais que o censo queria saber sobre a população dos Estados Unidos. Daí a utilização do termo *hollerite* para designar contracheque ou comprovante de pagamento dos trabalhadores.  
Fonte: [http://www.di.ufpb.br/raimundo/Revolucao\\_dos\\_Computadores/Histpage5.htm](http://www.di.ufpb.br/raimundo/Revolucao_dos_Computadores/Histpage5.htm).

nas escolas e universidades. Para tanto, foi criada a *International Association for Statistical Education* (IASE), com o objetivo de implementar e desenvolver mundialmente a Educação Estatística. Uma de suas ações foi a organização da *International Conference on Teaching Statistics* (ICOTS), cujo maior objetivo tem sido promover a oportunidade para que educadores estatísticos possam trocar informações, ideias, experiências e discutam as mais recentes inovações no campo do ensino da estatística. A primeira conferência ocorreu em 1982, na Inglaterra, e desde então vem sendo realizada de quatro em quatro anos. A ICOTS 7 foi realizada no Brasil, em Salvador, no período de 02 a 07 de julho de 2006, organizado pela Associação Brasileira de Estatística (ABE) e IASE (PAMPLONA e CARVALHO, 20--).

No Brasil, a primeira disciplina de estatística foi criada em 1863 na Escola Central, sucessora da Escola Militar, do Rio de Janeiro, com o nome de *Economia Política, Estatística e Princípios de Direito Administrativo*. Na Escola Politécnica em 1896, incluíam aplicações de *tábuas de mortalidade*<sup>14</sup>, problemas mais complicados de juros compostos, cálculos das sociedades denominados *Tontinas* e seguros de vida e propriedade, que faziam parte do currículo do curso de bacharel em Ciências Físicas e Matemáticas. Em 1925, na mesma escola, foi introduzida a Estatística Matemática. Em 1972, tem-se o primeiro vestibular para o curso de bacharel em Estatística, na Universidade Estadual do Rio de Janeiro (UERJ) e, no mesmo ano, na Universidade Estadual de São Paulo (USP) (PAMPLONA e CARVALHO, 20--). Na Universidade Federal do Paraná o curso de Estatística teve seu primeiro vestibular em 1974, conforme informações do site da universidade.

Atualmente, o Brasil conta com 38 cursos de graduação em Estatística reconhecidos pelo MEC (BRASIL, 2014a) e 17 cursos de pós-graduação *stricto sensu*, 10 de mestrado e 7 de doutorado, também reconhecidos (BRASIL, 2014b). Mas, a procura por esse tipo de curso no país ainda é muito baixa, apesar das vagas que sobram no mercado de trabalho e das excelentes perspectivas de carreira e salário. E ainda é insuficiente o número de profissionais formados para atender à demanda. Segundo notícia veiculada no site do Instituto de Ciências Matemáticas e

---

<sup>14</sup>A tábua de mortalidade, também conhecida como tábua de vida ou tábua de sobrevivência, é um instrumento que permite medir as probabilidades de sobrevivência e morte de uma população em função da idade, em um determinado momento ou período do tempo (ORTEGA, 1987, p.5). Disponível em: [http://www.esns.org.br/rbrs/arquivos/rbrs\\_16\\_2.pdf](http://www.esns.org.br/rbrs/arquivos/rbrs_16_2.pdf)

de Computação da USP, os estatísticos são disputados nas empresas, na academia e no governo, pois:

Em todos os setores, o emprego das metodologias estatísticas tem se tornado uma prática comum. “Hoje, o mercado absorve todos os bons estatísticos. Recebo constantemente pedidos para divulgar vagas de estágio aos alunos”, conta a professora Mariana Curi, coordenadora do curso noturno de bacharelado em Estatística do Instituto de Ciências Matemáticas e de Computação (ICMC) da USP São Carlos. “Existe, no Brasil, uma necessidade crescente por estatísticos e nós não damos conta de atender a demanda”, revela o professor do ICMC, Francisco Louzada. Ele realizou uma pesquisa, há dois anos, para avaliar a demanda por esses profissionais na área acadêmica e descobriu que, considerando-se todos os alunos que saem do doutorado, não é possível sequer preencher as vagas dos professores aposentados nos cursos de bacharelado em estatística nas universidades federais e estaduais brasileiras. Na visão do professor, essa situação deve ter se mantido desde então (USP, 2013).

A importância do estudo dessa ferramenta para análise e conhecimento da área social é reconhecida nos anos de 1980-90, pelo Ministério da Educação e do Desporto que, juntamente com as secretarias Estaduais de Educação, criou os *Parâmetros Curriculares Nacionais* (PCN). Uma das bases desses parâmetros é que o professor estabeleça ligações entre os conteúdos estudados com o cotidiano dos alunos, e que permita desenvolver suas competências e habilidades para o exercício da cidadania e um melhor convívio social (VARGAS, 2003). Para isso, foram introduzidos nos PCN, os conceitos elementares de estatística na Educação Básica, como conteúdo da matemática, com o nome de *Tratamento da Informação*. (BRASIL, 1998c), com finalidade de:

[...] fazer com que o aluno venha a construir procedimentos para coletar, organizar, comunicar dados, utilizando tabelas, gráficos e representação que aparecem frequentemente em seu dia-a-dia. Além disso, calcular algumas medidas estatísticas como média, mediana e moda com o objetivo de fornecer novos elementos para interpretar dados estatísticos. (BRASIL, 1998c, p.52).

Os PCN sugerem que os conteúdos do bloco *Tratamento da Informação* possam ser trabalhados em projetos interdisciplinares, integrado em outras áreas como História, Geografia, Biologia, Física, Química, etc., incentivando os alunos a observar fenômenos, reunir e analisar dados, especular hipóteses, interpretar gráficos, tabelas e medidas, de forma crítica (VARGAS, 2013).

### 2.3. POPULAÇÃO E AMOSTRA

Para utilizar a Estatística, primeiramente, deve-se conhecer o que é uma *população* e uma *amostra*. A *população*, como define MORAES (2005, p. 15), “pode ser considerada como uma coleção de Unidades Individuais”, isto é, pessoas ou resultados experimentais, com uma ou mais características entre si, na qual “o investigador quer fazer inferências ou estimativas”. *Amostra* é um número concreto de elementos extraídos de uma população (TORRECEILLA e GARRIDO, 2012).

Habitualmente, em uma pesquisa é muito difícil obter dados de toda população, devido a alguns fatores, como custo, dificuldade de aplicar o questionário em determinadas regiões, etc. A estratégia é selecionar uma amostra representativa para, a partir desta, estimar como é a população (TORRECEILLA e GARRIDO, 2012). Reis e outros conceituam que a base da inferência estatística consiste “na possibilidade de se tomarem decisões sobre os parâmetros de uma população, sem que seja necessário proceder a um recenseamento de toda a população” (REIS et. al., 2003 p. 21). Quando uma amostra é rigorosamente selecionada, Gil informa que:

[...] os resultados obtidos tendem a aproximar-se bastante dos auferidos se todos os elementos do universo fossem pesquisados. E, com o auxílio de procedimentos estatísticos, torna-se possível até mesmo calcular a margem de segurança dos resultados obtidos (GIL, 2010, p. 109).

Parâmetro é um dado necessário para analisar ou valorizar uma situação. A partir do parâmetro definido, pode-se interpretar determinada circunstância ou obter uma perspectiva da mesma. Na Estatística, é uma função definida sobre os valores numéricos de uma população, ou seja, é uma constante (valor fixo), geralmente desconhecido, que descreve características de uma população. Usualmente representados por caracteres gregos:  $\mu$ ,  $\rho$ ,  $\sigma^2$ ,  $\sigma$  etc. (FARIAS, 2008). Por exemplo, sendo a variável aleatória a idade dos professores de determinada rede, o número de elementos, no caso, corresponde ao número total de professores da rede, este é um parâmetro desta população. Desse total, posso calcular a média populacional de 35 anos, que também é um parâmetro da população. Note-se que isso é diferente da média amostral, ou seja, é diferente do valor que podemos obter tomando alguns professores e calculando a média de suas idades. Ou seja, a priori o parâmetro raramente é conhecido. Porém, em geral, é muito caro ou demorado obter dados da população inteira, por isso, usam-se amostras. Assim, de modo geral, os

caracteres usados para diferenciar as características da amostra das populacionais são os seguintes, como mostra o Quadro 2:

Quadro 2: CARACTERÍSTICAS PARA DIFERENCIAR POPULAÇÃO E AMOSTRA

| Característica      | Parâmetro (população) | Estatística (amostra) |
|---------------------|-----------------------|-----------------------|
| Número de elementos | N                     | n                     |
| Média               | $\mu$                 | $\bar{X}$             |
| Desvio Padrão       | $\sigma$              | S                     |

Fonte: adaptado de Farias, 2008.

O processo para determinar uma amostra, pode ser feito através de dois tipos de amostragem: amostragem probabilística e amostragem não probabilística. Neste trabalho, abordaremos apenas a probabilística, considerando a complexidade do tema e o recorte da pesquisa em foco.

*A amostragem probabilística é constituída por amostras nas quais todos os elementos de uma população possuem a mesma chance de participar da amostra, ou seja, “cada unidade elementar é sorteada com igual probabilidade, individualmente, e com um único estágio e seleção aleatória” (BOLFARINE; BUSSAB, 2005, p.16).*

*Para ter um método de seleção aleatória, é necessário definir um procedimento que garanta que as diferentes unidades da população tenham probabilidades iguais de serem escolhidas. Para isso deve-se cumprir 4 critérios, de acordo com Chochran (1977): (1) condição de definir o conjunto de amostras distintas que o procedimento é capaz de selecionar; (2) cada amostra deve ter uma probabilidade conhecida de seleção; (3) as amostras, selecionadas por um processo aleatório, devem ter, cada uma, a mesma probabilidade de ser selecionada; (4) o método para calcular a estimativa deve conduzir a uma estimativa única para qualquer amostra.*

*A amostragem probabilística pode ser classificada em: amostragem aleatória simples; amostragem sistemática; amostragem por clusters; amostragem por múltiplos estágios e amostragem aleatória estratificada.*

*A amostragem aleatória simples estabelece que cada unidade elementar seja sorteada com igual probabilidade, uma única vez e de forma aleatória. Exemplo sorteio da mega sena (BOLFARINE; BUSSAB, 2005).*

A *amostragem sistemática* é uma técnica que apresenta os elementos já ordenados, não necessitando criar um sistema de referência. Por exemplo, obter uma amostra de 20 professores de uma escola que contém 100 professores. Divide 100 por 20, o resultado é 5. Com uma lista desses professores seleciona-se um número aleatório, (por exemplo, o número 2). Na sequência, portanto, a seleção da amostra começa pelo 2º professor da lista, depois o 7º, o 12º e assim por diante, selecionando de 5 em 5, até compor os 20 professores da amostra (BOLFARINE; BUSSAB, 2005).

*Amostragem por Clusters* consiste em uma técnica utilizada por agrupamentos, relativamente homogênea. O significado da palavra *cluster*, em português, é agrupamento. Assim, na estatística, refere-se ao processo de classificação de objetos em diferentes grupos, cada um dos quais deve conter os objetos semelhantes. Nesta técnica, a população total está dividida em grupos (bairros, escolas, hospitais, etc.) e seleciona-se uma amostra aleatória simples dos grupos. Uma vantagem da amostragem de cluster é reduzir o número total de entrevistas e custos, dada a necessidade desejada. Por exemplo, selecionada uma escola, faz-se uma amostragem aleatória simples para selecionar a quantidade de professores que se deseja para fazer a pesquisa (BOLFARINE; BUSSAB, 2005).

*Amostragem por múltiplos estágios* é um tipo de amostragem com sorteios em vários estágios, primeiro seleciona-se o conglomerado (escola) e, dentro do conglomerado, sorteia-se o indivíduo (professor). Exemplo, muito usado em populações, onde se sorteia a cidade, depois o bairro, quarteirão, domicílio e depois o morador (BOLFARINE; BUSSAB, 2005).

*Amostragem aleatória estratificada* consiste na ideia de que, quanto mais homogênea for a população, melhor será o resultado. Uma vantagem da estratificação é a diminuição da variância, tendo em vista a homogeneidade dentro de cada estrato. (BOLFARINE; BUSSAB, 2005). A lógica do processo é que:

[...] dispondo dos itens da população global, o que leva à necessidade de um menor tamanho de amostra. É possível perceber isto considerando um caso extremo: Suponha idênticos os itens em cada estrato. Em tal caso, basta uma única observação de cada subgrupo para dizer do seu comportamento. Dessa forma, quanto maior a semelhança entre os elementos de cada estrato, menor o tamanho da amostra necessária (OLIVEIRA; ALMEIDA; BARBOSA, 2012, p. 12).

Para elaborar os grupos deve ser considerado o "efeito de aglomeração", o qual é muitas vezes medido pelo "coeficiente de correlação intracluster". Em se tratando de *análise de cluster*, Reis (1997) afirma que:

[...] é um processo utilizado em estatística multivariada, com vistas a organizar um conjunto de indivíduos, para os quais é conhecida uma informação detalhada, em grupos relativamente homogêneos, ou seja, a análise de cluster procura agrupar elementos de dados, baseando-se na similaridade entre eles (REIS, 1997, p. 287).

O plano de amostragem internacional preparado para pesquisa TALIS utilizou-se da *amostragem probabilística estratificada* em dois estágios. No primeiro estágio, o estrato referiu-se à seleção das escolas (público alvo de interesse da pesquisa); e no segundo estágio, o estrato referiu-se aos docentes do ensino secundário básico (nível 2 ISCED) e, para a seleção desses docentes, utilizou-se a amostragem aleatória simples (OCDE, 2009).

Porém, conforme já mencionado anteriormente, como os docentes da mesma escola tendem a compartilhar opiniões e se comportar de maneira semelhante, foi utilizado o *coeficiente de correlação intracluster* (OCDE, 2009).

Dessa forma, na pesquisa TALIS foi utilizado esse tipo de análise para controlar a homogeneidade dos professores e conhecer os grupos relativamente homogêneos, ou seja, essa análise visou agrupar elementos de dados, baseando-se na similaridade entre eles. Porém, é preciso ter cautela na utilização de dados dessa natureza, tomando-os com ponto de partida para análises de maior profundidade – especialmente, porque na educação são as diferenças e as desigualdades que devem ser objeto das políticas públicas voltadas à justiça social e à igualdade (REIS, 1997).

Isto permite compreender que, por exemplo, no programa TALIS, devido à homogeneidade de opiniões entre professores da mesma escola, tenha sido utilizado este tipo de análise, pois, o efeito de aglomeração é determinante para compor o número final de elementos da amostra (SIPRAKI; SALES; PEREIRA, 2012).



## 2.4. ESTATÍSTICA, CONCEITOS E MÉTODOS

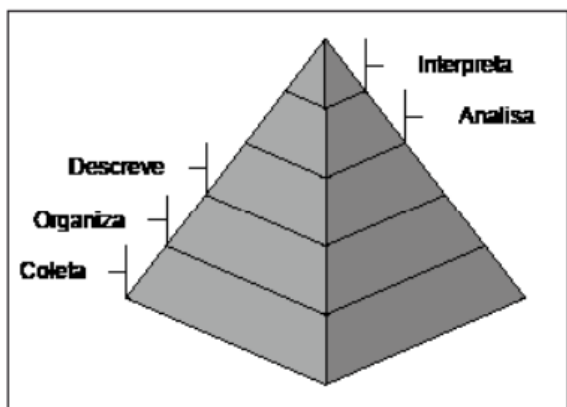
A Estatística, conforme já evidenciado, possui ferramentas importantes para trabalhar com dados empíricos e subsidiar tomadas de decisão que envolvem todas as áreas. Assim, a aplicação da estatística na área da educação tem se tornado cada vez mais utilizada. Abrange um conjunto de métodos e técnicas de pesquisa que inclui planejamento, coleta e organização de dados, descrição, correlações, inferência, processamento, análise, interpretação entre os dados e propagação de informações (MATTAR, 2001).

Esses métodos e técnicas possibilitam provar ou refutar hipóteses em um dado estudo. Nesse contexto, torna-se imperativo o auxílio de pessoas habilitadas em conhecimento estatístico para planejar, coletar, organizar e interpretar corretamente os dados necessários para realização de uma pesquisa. Em consenso com esse pensamento, Siegel (1975) afirma que a estatística,

[...] está interessada nos métodos científicos para coleta, organização, resumo, apresentação e análise de dados bem como na obtenção de conclusões válidas e na tomada de decisões razoáveis baseadas em tais análises (SIEGEL, 1975, p. 01).

Para tanto, se faz necessário conhecer o significado da estatística e suas técnicas, como pode ser observado na *pirâmide da definição* (Figura 2) elaborada por Medeiros (2007).

Figura 2: PIRÂMIDE DA DEFINIÇÃO



Fonte: Adaptado de CRESPO, 2002.

De acordo com a *pirâmide da definição* da estatística – que expressa na base que a estatística coleta, organiza e descreve e no topo, que analisa e interpreta os

dados – o mais importante é a interpretação dos dados, situada no patamar mais elevado. Porém, não é isso o que acontece em muitos estudos, ou seja, a maioria limita-se apenas à estatística descritiva, desconhecendo o que esta oferece de mais importante: “[...] o aspecto essencial da Estatística é o de proporcionar métodos inferenciais, que permitam conclusões que transcendem os dados obtidos inicialmente” (CRESPO, 2002, p. 13). É por meio de análises e interpretações que realmente se conhece a realidade do problema e através dessas interpretações se alcançam as conclusões necessárias, para não limitar-se aos *achismos* inconclusos e precários (MEDEIROS, 2007).

Quando se pretende realizar um estudo estatístico completo em determinada população ou em amostra de uma população, deverão ser obedecidas várias fases do método estatístico em geral, até chegar aos resultados finais procurados. Castanheira (2010) indica as principais fases que são: (1) *Definição do problema*, consiste em definir com clareza o objeto de estudo e qual é o objetivo a ser alcançado; (2) *Delimitação do problema*, que consiste em saber o que se pretende pesquisar e saber onde será realizada a pesquisa, ou seja, em que local, o tipo de pessoas (ou coisas), os dias, as horas, etc.; (3) *Planejamento para obtenção dos dados*, que consiste em responder: Como se irá resolver o problema? Quais dados serão necessários? Como obter esses dados? Esta fase pode ser desenvolvida por meio de observações ou, na maioria das vezes, por um questionário; também é importante estabelecer o cronograma das atividades, bem como o tamanho da população ou da amostra e quanto se pretende gastar para a realização da pesquisa; (4) *Coleta de dados*, consiste em obter os dados, através das observações ou do questionário aplicado. É a fase mais importante da pesquisa, pois, se a forma utilizada não atender às exigências metodológicas, terá ocorrido perda de tempo e de dinheiro; (5) *Apuração dos dados*, consiste em efetuar as críticas dos dados, ou seja o descarte dos dados que foram fornecidos de forma equivocada, ou seja, questionários mal respondidos ou mal aplicados; e resumir os dados através de sua contagem, separação por tipo de resposta e agrupamento de dados semelhantes, é o que se denomina de tabulação dos dados; (6) *Apresentação dos dados*, consiste em apresentar os dados em forma de tabelas ou em gráficos; (7) *Análise dos dados*, consiste em tirar as conclusões que auxiliem na solução do problema, está ligada ao cálculo de medidas que permitirão descrever, com

detalhes, o fenômeno que está sendo estudado; (8) *Interpretação dos dados*, consiste na interpretação, acompanhada de certo grau de incerteza, pois não se pode garantir *cem por cento* que os resultados obtidos numa amostra sejam totalmente verdadeiros para toda a população da qual a amostra pertence, portanto, o cuidado para concluir um estudo estatístico é de fundamental importância (CASTANHEIRA, 2010).

Conforme explicado por meio da pirâmide de definição de Medeiros (2007), os três primeiros degraus – coleta, organiza e descreve – referem-se à estatística descritiva e os dois últimos – analisa e interpreta – à estatística inferencial ou indutiva.

## 2.5. A ESTATÍSTICA DESCRITIVA

A estatística descritiva compreende conceitos básicos para determinar a forma que o estudo deve ser seguido por meio de levantamentos de dados de uma análise numérica. É através desse estudo que se observam possíveis tendências, erros de digitação, desenvolvem-se tabelas e gráficos e identifica-se o tipo da variável (numérica, ordinal, intervalar ou de razão) com a qual se irá trabalhar. Para Reis (2008) a estatística descritiva consiste “na escolha, apresentação, análise e interpretação de dados numéricos através da criação de instrumentos adequados: quadros, gráficos e indicadores numéricos” (REIS, 2008, p. 15).

Mas, antes de apresentar os conceitos e finalidades de uma estatística descritiva, é importante conhecer a natureza dos métodos estatísticos. Para tanto, Castanheira informa que:

São métodos para o tratamento de dados numéricos e referem-se a dados coletados, cujo destino é permitir que os estatísticos cheguem a conclusões sobre o que está sendo estudado (pessoas ou coisas) (CASTANHEIRA, 2010 p.15).

Essa natureza provém da fonte de observação que pode ser constituída pela população (universo) ou pela amostra. Reis (2008) define *população* como um conjunto de pessoas ou objetos que apresentam características em comum, que pode ser finita ou infinita. *Amostra* é uma parte da população a ser estudada (REIS 2008). Castanheira (2010) esclarece que o uso da amostra torna-se importante,

quando a população é muito grande e os custos muito elevados, o que determina a necessidade de selecionar uma parte finita dessa população (CASTANHEIRA, 2010).

As medidas descritivas são conhecidas, como aponta Reis (2008), por *parâmetros*, quando os dados provêm de uma população e *estatísticas*, quando provêm de uma amostra. Estes tipos de medida têm a função de sumariar os dados em um único valor que, para tanto, devem obedecer algumas propriedades para serem boas medidas no estudo em questão. Reis definiu que as medidas descritivas devem satisfazer os seguintes critérios:

- (i) Objetividade;
- (ii) Dependência de todas as observações;
- (iii) Significado bem preciso para a sua interpretação;
- (iv) Facilidade de cálculo;
- (v) Pouca variabilidade às flutuações de amostragem; e
- (vi) Facilidade de manejo no cálculo algébrico (REIS 2008, p.63).

Além disso, em estudos com coleta de dados amostrais, que representam uma população, as amostras estão associadas a uma margem de erro, sendo, portanto, de fundamental importância calcular suas estimativas. De acordo com Bolfarine e Bussab (2005, p.28), considera-se como *erro de estimação* aquele desvio devido apenas ao processo amostral e não relacionado a problemas de mensuração e obtenção das informações.

### 2.5.1 Conceito e Finalidades

A estatística trabalha com bancos de dados, ou seja, depois da coleta de dados, cria-se um banco de informações onde ficam armazenados todos os dados coletados. Importante ressaltar que a estatística descritiva não informa o resultado, apenas aponta erros de digitação e tendências dos acontecimentos. Para tomadas de decisão, faz-se necessário a estatística inferencial, que irá trabalhar com as análises dos dados preparados pela estatística descritiva, inferindo se a hipótese testada deve ser rejeitada ou não.

Nesse sentido, faz-se indispensável uma indagação: A educação usa a inferência para tomadas de decisão ou as pesquisas se dão por concluídas na fase descritiva? Muitas vezes, a forma como as questões são formuladas não permitem chegar à inferência.

A base, para qualquer análise de dados, consiste em estabelecer o tipo de variável adequada ao problema investigado, porque é através dela que é escolhido o método de análise mais apropriado. É nesse momento que a estatística descritiva começa o seu trabalho, categorizando (tabulando) o tipo de cada variável. Reis esclarece que *variável* é:

Um símbolo que representa determinada característica de uma população ou amostra de uma população. [...] Tal como uma característica, uma variável pode ser qualitativa ou quantitativa, discreta ou contínua. Toda a análise que fizer [...] irá distinguir entre variável discreta e contínua. Esta distinção torna-se absolutamente necessária por não se poderem aplicar os mesmos instrumentos estatísticos aos dois tipos de variáveis (REIS, 2008 p. 46).

Crespo (2004) salienta que, no conjunto de resultados possíveis de um fenômeno, a variável pode ser *qualitativa* quando expressa determinados atributos, como masculino e feminino, etc.; ou *quantitativa*, quando expressa números, como salários, idade, etc.

As variáveis são chamadas de *discretas*, quando expressas em números inteiros. Entre as qualitativas, a discriminação de sexo e cor do cabelo, por exemplo; e, entre as quantitativas, a ordem de classificação e grau de instrução, por exemplo.

As variáveis quantitativas também podem se *contínuas* quando podem assumir infinitos valores em um intervalo de números reais. Por exemplo: a altura de um indivíduo pode ser medida por meio de números decimais como 1,65m, 1,652m ou 1,6522m, conforme a precisão desejável. Assim:

Uma variável quantitativa que pode assumir, teoricamente, qualquer valor entre dois limites recebe o nome de variável contínua; uma variável que só pode assumir valores pertencentes a um conjunto enumerável recebe o nome de variável discreta (CRESPO 2002 p. 19).

Mas, na estatística descritiva, o mais importante é a *distribuição de frequência* e Castanheira (2010, p. 46) a define como “uma série estatística específica, que os dados estão dispostos em classes, com suas respectivas frequências absolutas”. Ou seja, os dados estão arrumados em números distribuídos em tabelas e, através destas, podem ser criados gráficos, identificadas *medidas de tendência central* e de *dispersão* que serão tratadas em detalhe mais adiante, após a conceituação sobre as escalas de medida.

### 2.5.2. Escalas de Medida

Para conhecer as *Escala de Medida*, é importante compreender primeiro a distinção entre mensuração e medida. *Mensuração* é o processo do qual resulta a medida, ou seja, atribuir certas grandezas (valores) aos dados obtidos, e *medida* é o valor (número) resultante da mensuração, portanto, medir é atribuir um número e a esse número atribui-se um nível ou escala, que é utilizada para a análise estatística (CRESPO 2004).

A base para qualquer análise de dados consiste em estabelecer o tipo de medida que deve ser utilizado, porque é através desta que é escolhido o método de análise mais apropriado. Assim como existem características ou atributos que podem ser medidos, em contrapartida existem os que não podem, sendo necessário criar instrumentos mais apropriados. Reis (1997) afirma que existem duas qualidades básicas para que um indicador represente o conceito teórico para o qual foi criado e que se propõe a medir que são: *validade* e *fiabilidade*. Sobre essas qualidades, a autora esclarece que:

Um instrumento de medida de determinado conceito abstrato é válido se conseguir efetivamente medir aquilo para que foi desenvolvido e é fiável se medir os mesmos resultados em experiências repetidas em idênticas condições. Por exemplo, um teste feito para aferir os conhecimentos absorvidos por um grupo de alunos ao longo de um ano letivo é válido se conseguir efetivamente distinguir os alunos com maior aprendizagem dos que nada conseguiram absorver ao longo do ano (REIS, 1997 p. 29).

Para elaborar instrumentos de coleta de dados são utilizadas escalas de medida. Cada escala de medida possui seu próprio conjunto de pressupostos, referentes à correspondência com números e à realização das várias operações matemáticas com esses números.

Entre os vários tipos de escala que podem ser utilizados, a estatística propõe: *nominal*, somente para identificação e definição de números; *ordinal*, para ordenação numérica; *intervalar*, para comparação de intervalos; e *de razão*, para comparação de medidas absolutas e de proporções (MATTAR, 2001). Assim, é necessário conhecer as possibilidades de cada uma das escalas. Para cada uma delas, existem possibilidades e limitações. Nesse caso, a escolha está condicionada a estes fatores. Segundo Oliveira:

[...] os pesquisadores devem estar atentos para identificar quais as características do constructo estão sendo medidas e quais são as propriedades do sistema numérico que se referem ao constructo para, a partir daí, escolher a técnica que melhor se adapta ao problema de pesquisa. Em suma, é possível dizer que a escolha de um formato apropriado para a pesquisa deve levar em consideração a natureza da variável a ser medida, a habilidade dos respondentes de fazer julgamentos e os tipos de análise a serem desenvolvidos (OLIVEIRA, 2001, p.03).

Em uma *escala nominal*, os números servem apenas para categorizar dados sobre pessoas, objetos, etc. Ou seja, é um tipo de escala menos sofisticado, na qual os números servem como legendas para identificar diferentes categorias. Este rótulo pode ser numérico, porém esses números não são utilizados para contagem do número de respostas, mas para identificar cada categoria. Exemplo: sexo (feminino [1] e masculino [2]), cidades (1, 2, 3, 4, ...), graus de instrução, estado civil, cor dos olhos, etc. (OLIVEIRA, 2001).

As *escalas ordinais* distribuem-se de certa ordem, que pode ser crescente ou decrescente, permitindo estabelecer diferenciações, ou seja, podem nomear identificar e categorizar objetos, pessoas ou fatos. Nesse caso, utiliza-se para estabelecer uma classificação: 1º lugar, último lugar, nível social, nível salarial, e escalas usadas nas medidas de opinião, como por exemplo, a Escala Likert, que é muito utilizada nos questionários da pesquisa TALIS, para identificar as crenças dos professores.

A Escala Likert surgiu através de um relatório publicado em 1932, cujo autor “Rensis Likert”, explica o método que aborda um tipo de escala de resposta psicométrica. É a mais utilizada em questionários e pesquisas de opinião, onde os entrevistados respondem perguntas baseado em escalas através de níveis de discordância ou concordância (LIMA et. al. 2012). Desse modo:

Em geral são utilizadas na escala de Likert quatro ou cinco categorias ordinais. Como ilustração pode ser citada, para quatro categorias, 0 - nada importante, 1 - pouco importante, 2 - importante e 3 - muito importante, e para cinco categorias, 0 - muito baixo, 1 - baixo, 2 - médio, 3 - alto e 4 - muito alto. [...] Uma das grandes preocupações em qualquer pesquisa, em particular, aquelas onde o elemento humano é a unidade a ser pesquisada ou o fornecedor das informações investigadas com base na sua percepção, é o fiel registro dessas informações, isto é, o que se deseja registrar é a opinião do entrevistado que retrate a realidade do fenômeno estudado (ALEXANDRE et. Al. 2003, p. 1-2).

Os mesmos autores informam que, na escala de 0 a 4 não existe uma categoria central e isso pode induzir dos respondentes a uma determinada

tendência. Incluir a opção “não sei” é uma sugestão para evitar esse problema na construção da escala. Na escala de 0 a 5, o problema pode aparecer com a categoria do meio, representando uma indecisão. Podem-se utilizar outras escalas, que variam “de quatro a onze categorias, mas as escalas de quatro e de cinco são as mais populares” (ALEXANDRE et. al., 2003, p. 3).

A *escala intervalar* é uma forma quantitativa de registrar um fenômeno, medindo a sua intensidade específica, ou seja, “além da classificação e ordenação dos seus valores, permitem o cálculo de distâncias ou diferenças entre observações” (REIS, 1997, p. 31). Portanto, este tipo de escala possui uma distância entre si com valores reais em relação à determinada característica. MORAES (2005 p. 6) informa que nessa escala “a diferenciação dos indivíduos ou das observações assume um valor quantitativo constante”, e que esses valores “envolvem classificação, grandeza e unidades de tamanhos idênticos”. Um exemplo clássico são as escalas de temperatura, onde não se pode assumir um ponto zero como ausência de temperatura, ou dizer que a temperatura X é o dobro da temperatura Y. (MORAES, 2005).

Finalmente, as *escalas de razão* possuem, além das mesmas propriedades das escalas de intervalo, a vantagem de medir a magnitude absoluta e quanto as variáveis (pessoas ou objetos) estão distantes entre si em relação à determinada característica (OLIVEIRA, 2001). É importante observar que, nessa escala, o valor “2”, indica que realmente é o dobro do valor “1”, por exemplo, o que não acontece necessariamente nas outras escalas, (MORAES, 2005). Todavia,

O uso de diferentes tipos de escalas não constitui problema, desde que seus referenciais apresentem pontos comuns que os tornem equivalentes, o que nem sempre ocorre. Assim, os grandes referenciais são quase sempre a média, o desvio padrão e o chamado escore “z”<sup>15</sup>, que expressa a relação da diferença entre o escore obtido e a média do grupo em termos de desvio padrão (VIANNA 2003, p. 63-64).

Assim sendo, quando se têm valores de certa característica, é fácil constatar que os dados normalmente não se distribuem uniformemente, havendo certo agrupamento. Pode-se, portanto, estudar os valores numéricos que determinam a distribuição dos dados, procurando o ponto onde está a maior concentração de

<sup>15</sup> Escore “z” ou distribuição normal de probabilidades é uma distribuição contínua, simétrica em relação à média, representada por uma curva em forma de sino, conhecida como “Curva de Gauss”. Qualquer conjunto de valores X, normalmente distribuídos, pode ser convertido em valores normais z (CASTANHEIRA, 2010).



valores individuais. De um modo geral, um conjunto de dados pode ocupar uma posição específica dentro de uma distribuição. Essas medidas que "posicionam" o dado (ou o grupo de dados) dentro de uma distribuição são chamadas de *medidas de tendência central*, que englobam média, mediana e a moda – que representam elementos pelos seus valores médios, em torno dos quais se concentram os dados – e as *medidas de dispersão*, que envolvem o desvio padrão e o coeficiente de variação, que serão vistos no próximo tópico (REIS 2008).

### 2.5.3. Medidas de Tendência Central e Dispersão

As medidas de tendência central são indicadores que permitem que se tenham uma ideia ou um resumo, de um determinado conjunto de dados, mostrando apenas um valor específico. E a dispersão são valores que estão em torno da medida de tendência central, identificando a distância de cada valor.

Assim, nesse tópico será apresentado o significado da média aritmética, a mediana e a moda como medida de tendência central e o desvio padrão e o coeficiente de variação e da amplitude total como medida de dispersão.

A *média* é o valor de um conjunto de dados que se localiza em um ponto central. Por essa razão, medidas com essa tendência são também denominadas medidas de tendência central. Vários tipos de média podem ser definidos, sendo as mais comuns a média aritmética simples e a média aritmética ponderada (CRESPO 2004).

A *média aritmética simples* é a soma dos resultados obtidos divididos pelo número que expressa sua quantidade. Em um conjunto com vários dados ( $x_1, x_2, x_3, x_4, \dots$ ),  $a = (x_1 + x_2 + x_3 + x_4 + \dots) / n$  ou:

$$media = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} \quad \text{ou} \quad \bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (1)$$

Fonte: CRESPO, 2004.

Na fórmula explicitada acima, temos:

$\sum x$  = soma total dos valores da variável;

$x_i$  = valores individuais observados;

$n$  = dimensão da amostra.

De tal modo, o cálculo da média aritmética simples é calculado, por exemplo, a partir do conjunto de valores de uma amostra de 5 observações – 10; 2; 9; 6; 8 –  $M = (10 + 2 + 9 + 6 + 8) / 5 = 7$ .

No exemplo citado, o valor 7 corresponde a um "ponto de equilíbrio" (valor em torno do qual os dados se distribuem). Só se deve arredondar a média quando ela representar variáveis quantitativas discretas como, por exemplo, idade, número de filhos etc., as quais não podem ser expressas com números fracionados (REIS 2008).

A *média aritmética ponderada* é utilizada quando os dados estão distribuídos em intervalos de classes (distribuição de frequência). Para esse cálculo, usa-se a média aritmética dos valores ( $x_1, x_2, x_3, x_4, \dots, x_n$ ), ponderados pelas frequências absolutas dos intervalos de classes ( $f_1, f_2, f_3, f_4, \dots, f_n$ ). Isso significa dizer que cada grandeza envolvida no cálculo da média aritmética simples, tem diferente importância, ou seja, incidu um número diferente de vezes durante a coleta de dados (CASTANHEIRA, 2010). Esta média é calculada a partir da seguinte fórmula (ver Castanheira, 2010):

$$\bar{X} = \frac{\sum_{i=1}^n (x_i \cdot f_i)}{N} \quad (2)$$

Fonte: Castanheira 2010

Nesta fórmula, temos:

$\Sigma$  = soma total dos valores da variável

$x_i$  = valores individuais observados;

$n$  = dimensão da amostra

$f_i$  = frequência

$x_i f_i$  = número de observações multiplicado pela frequência = ponderação

O modo mais prático para calcular uma média ponderada é adicionar uma coluna na *Tabela de Distribuição de Frequência*. Nesta, conforme exemplo de Castanheira (2004), temos uma coluna com a idade das pessoas observadas e outra a "frequência", conforme exemplifica a Tabela 1:

Tabela 1: IDADE DE UM GRUPO DE PESSOAS

| Idade | Frequência (f) |
|-------|----------------|
| 4     | 4              |
| 5     | 6              |
| 6     | 6              |
| 7     | 4              |

Fonte: CASTANHEIRA, 2010.

No exemplo citado, a aplicação da fórmula consiste no seguinte cálculo:

$$\bar{x} = \frac{4 \cdot 4 + 5 \cdot 6 + 6 \cdot 6 + 7 \cdot 4}{20} = 5,5$$

A idade média desse grupo é de 6 anos (resultante do arredondamento de 5,5). A ponderação simplifica o cálculo que poderia ser obtido pela média aritmética simples, com a soma das frequências de cada uma das idades observadas somada e dividida pelo número de frequência total.

Outra forma de ponderação pode ser utilizada com a atribuição de pesos. Como por exemplo, para calcular a média de notas dos alunos, onde cada instrumento de avaliação possui um valor diferenciado.

A média aritmética é uma das medidas mais usadas em análises estatísticas, porém, não é uma medida satisfatória para todos os dados estatísticos, “quando ela é afetada pelos grandes valores” (CRESPO 2002 p. 90). Por exemplo, no caso da idade média de uma turma com 10 alunos, poderíamos ter  $12 + 11 + 13 + 12 + 11 + 13 + 13 + 12 + 12 + 40 / 10 = 15$  anos. Ou seja, o aluno com 40 anos de idade afeta a média de tal forma, que o resultado é insatisfatório.

Outra forma de analisar o número de alunos por turma é através da *mediana*, medida esta definida por Crespo como:

o número que se encontra no centro de uma série de números, estando estes dispostos segundo uma ordem. Em outras palavras, a mediana de um conjunto de valores, ordenados segundo uma ordem de grandeza, é o valor situado de tal forma no conjunto que o separa em dois subconjuntos de mesmo número de elementos. (CRESPO, 2004, p. 89).

Castanheira (2010), por sua vez, a define como o valor que ocupa a posição central em um conjunto de dados colocados em ordem crescente ou decrescente, ou seja, 50% para menores valores de um lado e 50% de maiores valores de outro

lado. Para calcular a mediana, faz-se necessário observar a quantidade de dados e, se o número total da população ou amostra for par, somam-se os dois valores centrais e divide-se por dois, o resultado é a mediana. Se for ímpar, a mediana corresponde ao valor central.

No exemplo citado anteriormente, para observações com número par, em um conjunto de dados com 10 observações, cujos valores apresentados em ordem crescente são 11 +11+ 12 + 12 +12 +12 + 13 +13 +13 + 40 anos, a mediana é o valor correspondente à soma dos 2 valores centrais dividido por dois. O valor da mediana para esse caso é:

$$Md = \frac{12+12}{2} = 12 \text{ anos}$$

Se acrescentarmos dois alunos de 13 anos e outro com 14 anos no mesmo exemplo, para observações com número ímpar, teremos um conjunto de dados com 13 observações, cujos valores em ordem crescente são 11 +11+ 12 + 12 +12 +12 + 13 + 13 + 13 +13 +13 + 14 + 40 anos. A mediana nesse caso será o valor correspondente à posição central, que é  $Md = 13$  anos.

O interessante é observar que o uso da mediana não afeta o valor central ao contrário da média aritmética simples.

Para decidir qual das duas medidas deve ser aplicada, fica evidente que, quando os dados estiverem dispersos, ou seja, com alguns valores muito grandes ou muito pequenos, destoando dos demais, é melhor utilizar a mediana.

Essa medida é pouco utilizada na área educacional, assim como se pode perceber nas análises da pesquisa TALIS ou mesmo em outras pesquisas de larga escala citadas anteriormente – como SAEB, ENEM, etc. – nas quais sequer foi testada.

A *moda* é uma das medidas de fácil resolução e para encontrá-la, basta verificar o número que mais se repete nas observações. Crespo a define como:

o valor que ocorre com maior frequência, isto é, o valor mais comum. É assim que podemos dizer que “o salário modal dos empregados de uma indústria é o salário mais comum, isto é, o salário recebido pelo maior número de empregados dessa indústria”. (CRESPO, 2002, p. 94).

A moda possui algumas vantagens em relação às outras já citadas. É fácil de calcular e não é afetada por valores extremos, haja visto que seu valor é o que mais se repete, e a desvantagem porque não pode ser definida com rigor e o seu valor exato é, muitas vezes, incerto (REIS, 2008).

Medir somente os valores centrais em um estudo, não é suficiente para caracterizar de forma adequada uma variável e, por essa razão, é de suma importância analisar a *medida de dispersão* que está em sua volta. Reis define que esta medida:

[...] serve para verificarmos a representatividade das medidas de localização [medidas de tendência central], pois é muito comum encontrarmos variáveis que apesar, de terem a mesma média, são compostas por valores bem distintos (REIS, 2008 p. 97).

Para avaliar o grau de variabilidade (dispersão ou afastamento) de um conjunto de números, Castanheira (2010) explica que se utilizam as *medidas de dispersão*, que permitem conhecer com mais detalhe os fenômenos em questão. Crespo define medidas de dispersão como um conjunto referente à maior ou menor variabilidade dos valores de uma variável em torno de uma medida de tendência central, ou seja, permitem avaliar se há ou não homogeneidade nos dados. A estatística recorre a essas medidas quando deseja qualificar os valores de uma variável. Para este estudo, será analisado o desvio padrão e o coeficiente de variação (CRESPO 2004).

Considerando a importância de calcular o desvio em torno da média, torna-se útil estudar o *desvio padrão*, cujo significado é a medida da variação em um determinado conjunto de dados. Assim, quanto maior for o desvio padrão, maior será a heterogeneidade entre os valores que estão sendo analisados. Isso significa, portanto, que quanto mais heterogêneos são os dados, maior será a variação entre os valores. A fórmula para o cálculo do desvio padrão é:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} \quad (3)$$

Em que,

- $\bar{x}$  = média aritmética da série
- $n$  = tamanho da população
- $x_i$  = dado da série
- $\sigma$  = desvio-padrão

Obs.:  $(x_i - \bar{x})$  = dispersão de cada um dos dados da série em relação à média.

Fonte: CRESPO, 2004.

O *coeficiente de variação* é uma forma de caracterizar com maior rigor a dispersão dos conjuntos de dados e possui a função de compreender os termos relativos em forma de percentual, ao contrario do desvio padrão que possui a função de compreender os termos absolutos (REIS, 2008). Ou seja, é uma medida de dispersão empregada para estimar a precisão de experimentos e representa o desvio padrão como porcentagem de dispersão em torno da média. Como duas distribuições podem ter médias diferentes, o desvio-padrão dessas duas distribuições não é comparável. A solução é usar o coeficiente de variação, que é igual ao desvio-padrão dividido pela média. A fórmula para o coeficiente de variação é definida como:

$$CV = \frac{S}{X} \cdot 100 \quad (4)$$

Fonte: REIS, 2008

Em que,

CV = Coeficiente de variação;

S = Desvio padrão

X = média aritmética

Sua principal qualidade é a capacidade de comparar resultados de diferentes médias que envolvem o mesmo valor do desvio padrão. Em igualdade de condições, é mais preciso o experimento com menor coeficiente de variação. Dessa forma, Mohallen e outros (2008, p. 450) classificam os coeficientes de variação como “baixos quando inferiores a 10%, médios entre 10 a 20%, altos entre 20 e 30% e muito altos se superiores a 30%”.

Assim, utiliza-se esse cálculo para medir o percentual de dispersão observado. O coeficiente de variação é muito útil para estudos na área da educação. Como exemplo, pode-se analisar, em uma pesquisa hipotética, a nota média de duas turmas de alunos matriculados no 6º ano do ensino fundamental, na qual o(Quadro 03) mostra as notas médias, os desvios padrão e os coeficientes de variação. Pode-se perceber que a turma do 6º ano B tem uma dispersão maior do que a do 6º ano A; no entanto, o desvio padrão é igual em ambas.

Quadro 3: PESQUISA HIPOTÉTICA PARA EXEMPLIFICAR O COEFICIENTE DE VARIAÇÃO

| Turma    | Nota média | Desvio padrão | Coeficiente de variação     |
|----------|------------|---------------|-----------------------------|
| 6º ano A | 40         | 4             | $\frac{4}{40} = 0,1$ ou 10% |
| 6º ano B | 5          | 4             | $\frac{4}{5} = 0,8$ ou 80%  |

Fonte: Elaboração do autor.

Quanto ao coeficiente de variação, tem-se respectivamente 0,1 e 0,8, ou seja, pode-se afirmar que, na turma do 6º ano A, os desvios relativos à nota média atingem em média 10%. Na turma do 6º ano B, porém, os desvios relativos à nota média atingem 80%. De acordo com a classificação de MOHALLEN, (2008), as notas dos alunos do 6º ano A apresentam baixa dispersão em relação à nota média, enquanto que as notas dos alunos do 6º ano B apresentam alta dispersão em relação à nota média.

*Amplitude total* é uma das mais simples medidas de dispersão, que pode ser definida com a subtração do limite máximo com o limite mínimo de um grupo de observações. Pode ser simbolizada pela letra “A” e é calculada da seguinte maneira:

$$A = X_{\text{máx}} - X_{\text{mín}} \quad (5)$$

Onde:

A: Valor da amplitude total;

$X_{\text{máx}}$ : Valor do limite máximo;

$X_{\text{mín}}$ : Valor do limite mínimo.

Fonte: Adaptado de REIS, 2008.

Um exemplo desta dispersão é, considerando que as notas de quatro bimestres do ano letivo de um aluno sejam 60, 60, 60 e 60; calcula-se a dispersão tomando o limite máximo (60) e subtraindo o limite mínimo (60), obtendo como resultado  $A = 0$ , ou seja, não há dispersão. Em outro exemplo sobre as notas bimestrais de outro aluno, sendo 30, 55, 80 e 80; o limite máximo (80) menos o limite mínimo (30), obtém como resultado  $A = 50$ , isto é, a dispersão é de 50 pontos.

Porém a amplitude não é uma medida completa de variação, porque o cálculo utiliza apenas as variações extremas (dado maior e menor do grupo), não avaliando

os valores intermediários, desse modo, seu valor tende a crescer de acordo com o aumento do número de observações (REIS, 2008).

Com a aplicação adequada das medidas de tendência central e das medidas de dispersão, os resultados serão, estatisticamente, mais precisos e as análises, mais coerentes com os fatos. No caso da educação, onde na maioria das pesquisas utiliza-se somente a estatística descritiva, o uso apropriado dessas medidas pode evidenciar tendências mais próximas à realidade. De modo geral, as pesquisas educacionais divulgam apenas as médias que, nem sempre são as medidas mais adequadas para compreender determinado problema, e o desvio padrão, que mostra o grau de homogeneidade ou heterogeneidade da referida média, é ignorado. Este é o caso da média obtida no PISA ou na Prova Brasil, por exemplo. Portanto, é de fundamental importância averiguar se a melhor solução para determinado estudo é o uso da média ou da mediana e, sobretudo, avaliar o grau de dispersão.

Muitos entendem que a estatística se resume à criação de tabelas e gráficos para analisar valores econômicos, resultados eleitorais ou levantamentos de dados quantitativos sobre pessoas e objetos (MORETTIN 1981), como ocorre, por vezes, na área da educação.

Mas, a estatística moderna com seus softwares de última geração não é responsável somente pela criação de dados descritivos, trabalha também com metodologias científicas mais complexas para responder perguntas e questionamentos, tais como: O fumante passivo pode vir a desenvolver câncer? Pode determinado medicamento reduzir o risco de um ataque cardíaco? Qual será o preço do ouro no final deste ano? O método de avaliação que o professor aplica em sala de aula é adequado? O número excessivo de alunos por turma atrapalha o desempenho docente? (UFFE, 2015). Para responder a essas perguntas e tomar decisões sobre a solução de problemas, é necessário reconhecer que todo estudo está passível de erros, significando, com isto, que se deve respondê-las de forma a minimizar o risco envolvido. Para tanto, a estatística é a ciência que permite extrair dos dados as informações necessárias com base em intervalos de confiança e redução da margem de erro, que a inferência estatística poderá responder.



## 2.6. ESTATÍSTICA INFERENCIAL

A *Estatística Inferencial* é uma parte da Estatística que compreende métodos e procedimentos para deduzir características de uma população, a partir de uma amostra (TORRECEILLA e GARRIDO, 2012). Como a *inferência estatística* possui um vasto campo de trabalho com diferentes aplicações e métodos que dependem de cada amostra estudada, neste trabalho, as explicações e conceitos se limitarão somente aos procedimentos que serão usados como exemplo para analisar a comparação da média de alunos por turma a partir do banco de dados da pesquisa TALIS.

Segundo Neves (2009), o objetivo da *inferência estatística* está na construção e desenvolvimento de métodos para explicar, através de uma amostra, conclusões precisas acerca de uma população estudada. A autora sustenta que, em um problema existem dois tipos de inferência, que se determinam por meio de *testes paramétricos* e *não paramétricos*. Os *testes paramétricos* baseiam-se em escalas intervalares e de razão, que obedecem a uma distribuição normal (onde os dados possuem certa simetria, com média e desvio padrão homogêneos); e os *testes não paramétricos* baseiam-se em dados ordinais e nominais ou quando a amostra não atende aos requisitos de uma distribuição normal (NEVES, 2009).

Dentre os testes paramétricos, o teste *t* de *Student*<sup>16</sup> é proposto para comparação de médias provenientes de duas amostras independentes. Estes tipos de comparação são os dois métodos mais utilizados para esta finalidade. Importante ressaltar que para esses dois métodos, habitualmente, usam-se a média como medida de tendência central e o desvio padrão como medida de dispersão. É através dessas medidas que os testes são realizados (NEVES, 2009).

Em todos os testes inferenciais, as decisões acerca das características da população baseiam-se em hipóteses que, de um modo geral, são afirmações suportadas pelas distribuições dos dados relativos às características da amostra

---

<sup>16</sup>Teste t de Student: A estatística t foi introduzida em 1908 por William Sealy Gosset, químico da cervejaria Guinness em Dublin, Irlanda ("*student*" era seu pseudônimo). Gosset havia sido contratado devido à política inovadora de CLaude Guinness de recrutar os melhores graduados de Oxford e Cambridge para os cargos de bioquímico e estatístico da indústria Guinness. Gosset desenvolveu o Teste t como um modo barato de monitorar a qualidade da cerveja tipo *stout*. Ele publicou o Teste t na revista acadêmica *Biometrika* em 1908, mas foi forçado a usar seu pseudônimo pelo seu empregador, que acreditava que o fato de usar estatística era um segredo industrial. De fato, a identidade de Gosset não foi reconhecida por seus colegas estatísticos (NEVES, 2009).

populacional que se pretendem testar, e que Castanheira, (2010 p. 218) define “como uma técnica que nos permite aceitar ou rejeitar a hipótese estatística, a partir dos dados da amostra dessa população”. Isso significa dizer que “a partir de um teste realizado com os dados de uma amostra, pode-se inferir sobre a população a que essa amostra pertence”(CASTANHEIRA, 2010 p. 218). Para tanto é necessário formular dois tipos de hipóteses: a *hipótese nula* e a *hipótese alternativa*.

A *hipótese nula*, designada por  $H_0$ , é uma hipótese tida como verdadeira até que se prove estatisticamente o contrário. A partir da impossibilidade de observar toda a população, o teste é baseado na informação a respeito da média e do desvio padrão da amostra em estudo. E a *hipótese alternativa*, designada por  $H_1$ , é a hipótese que se utiliza quando se comprova que a hipótese nula é falsa, isto é, que as médias das duas amostras do estudo diferem (MORAES, 2005).

TORRECILLA e GARRIDO (2012 p. 43) descreve matematicamente as hipóteses da seguinte forma:

$$H_0: \mu_1 = \mu_2 \quad (6)$$

$$H_1: \mu_1 \neq \mu_2$$

Em que,

$H_0$ : hipótese nula;

$H_1$ : hipótese alternativa;

$\mu_1$ : média da primeira população

$\mu_2$ : média da segunda população

Segundo Castanheira (2010), as *estimativas por intervalos*, também denominadas de *Intervalo de Confiança* (IC), corresponde a uma faixa de valores possíveis, obtidos a partir de observações de uma amostra, dentro da qual se estima que está o parâmetro que se deseja determinar. Essa estimativa permite diminuir a magnitude do erro amostral, ou seja, quanto menor o comprimento do intervalo, maior será a precisão do cálculo, dentro de um nível de confiança<sup>17</sup> que tenha uma

<sup>17</sup>Nível de Confiança: corresponde a  $1 - \alpha$  (probabilidade de aceitar a hipótese nula quando esta é verdadeira) e indica a probabilidade de decisão correta baseada na hipótese nula. O valor  $\alpha$ , é a probabilidade de que o intervalo encontrado NÃO contém o valor do parâmetro procurado, isto é, a probabilidade de que o valor do parâmetro em estudo tem na realidade um valor fora do intervalo construído. Os níveis de significância mais utilizados são 5%, 0,1%, 1% e 10%. De acordo com os padrões internacionais o nível de significância mais utilizado é 5%

Fonte: <http://www.lee.dante.br/pesquisa/amostragem/glossario.html>

chance de  $1 - \alpha$ , ou seja, é a probabilidade de que o intervalo encontrado contém o valor do parâmetro. O IC para a média é expresso pela fórmula:

$$I.C.(1-\alpha): \bar{X} - t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \quad (7)$$

Fonte: Adaptado de NEVES(2009)

Em que,

$1 - \alpha$ : Nível de confiança;

X: Média da amostra;

$t_{\alpha/2}$ : Grau de liberdade (habitualmente determinado por 1,96, quando o nível de confiança for 95%);

s: Desvio padrão da amostra;

n: Amostra;

$\mu$ : média populacional.

Na maioria dos softwares, a significância estatística é expressa pelo nível descritivo (valor p)<sup>18</sup>, a qual é a probabilidade que permite decidir sobre a validade ou não da hipótese nula. MORAES (2012) exemplifica algumas interpretações acerca do valor p:

- Se  $p \geq 0,1$ , não existe evidência contra a  $H_0$ , não é possível rejeitar a hipótese nula;
- Se  $p < 0,1$ , fraca evidência contra a  $H_0$ , que não permite certeza;
- Se  $p < 0,05$ , evidência significativa contra a  $H_0$ , é possível rejeitar a hipótese nula;
- Se  $p < 0,01$ , evidência altamente significativa contra a  $H_0$ , é possível rejeitar a hipótese nula;
- Se  $p < 0,001$ , evidência muito altamente significativa contra a  $H_0$ , é possível rejeitar a hipótese nula. (MORAES, 2012 p. 22).

O teste  $t$  de Student, que será utilizado para comparação de médias nesta pesquisa, é definido por D'Hainaut (1990 p. 192) como o "método [que] permite decidir se a diferença observada entre as médias de duas amostras podem se atribuir a uma causa sistemática ou se pode ser considerada como efeito de flutuações devidas ao acaso". Esse teste analisa se a média da amostra é diferente ou não da média populacional.

<sup>18</sup> Valor p: corresponde ao menor nível de significância que pode ser assumido para rejeitar a hipótese nula. Dizemos que há significância estatística quando o valor p é menor que o nível de significância adotado. Por exemplo, quando  $p=0.0001$  pode-se dizer que o resultado é bastante significativo, pois este valor é muito inferior aos níveis de significância usuais. Se  $p=0.048$  pode haver dúvida, pois, embora o valor seja inferior, ele está muito próximo ao nível usual de 5%. Fonte: <http://www.lee.dante.br/pesquisa/amostragem/glossario.html>

Há uma forma de representação matemática para o cálculo do teste t para duas amostras independentes. A apresentação das Hipóteses:

$$H_0: \mu_1 = \mu_2 \quad (8)$$

$$H_1: \mu_1 \neq \mu_2$$

Em que,

$H_0$ : as médias não diferem;

$H_1$ : as médias diferem.

E para representar o teste *t* Student para duas amostras independentes é:

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{s_1^2 \times (n_1 - 1) + s_2^2 \times (n_2 - 1)}{(n_1 + n_2 - 2)} \times \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad (9)$$

Fonte: adaptado de NEVES, (2009)

Em que,

*t*: Valor do teste;

$\bar{x}_1$ : média da amostra 1;

$\bar{x}_2$ : média da amostra 2;

$s_1^2$ : Variância (desvio padrão ao quadrado) da amostra 1;

$s_2^2$ : Variância (desvio padrão ao quadrado) da amostra 2;

$n_1$ : Tamanho da amostra 1;

$n_2$ : Tamanho da amostra 2;

Teoricamente, os testes paramétricos são os melhores e mais poderosos para testar hipóteses. Porém, na prática, nem sempre acontece dos dados estarem dentro dos pré-requisitos exigidos pelos testes. É muito comum encontrar questionários respondidos erradamente, faltando dados, amostras muito pequenas, respostas com valores absurdos (muito conhecidos na estatística como *outliers*)<sup>19</sup> e variáveis categoricamente mensuradas como nominais e ordinais. Para tanto, a

<sup>19</sup> *Outliers*: é caracterizado pela sua relação com as restantes observações que fazem parte da amostra. O seu distanciamento em relação a essas observações é fundamental para fazer a sua caracterização. Estas observações são também designadas por observações "anormais". Grande parte dos autores que estudam este fenómeno se referem a comentários de Bernoulli datados de 1777 [ ver por exemplo, Barnett e Lewis(1994)] , como sendo uma das primeiras e mais importantes referências a observações *outliers*. Esses comentários indicam que a prática de rejeitar tal tipo de observação era comum naquela altura (século XVIII). A discussão sobre os *outliers* centrava-se na justificação da rejeição daqueles valores. As opiniões não eram unânimes: uns defendiam a rejeição das observações "inconsistentes com as restantes", enquanto outros afirmavam que as observações nunca deviam ser rejeitadas simplesmente por parecerem inconsistentes com os restantes dados e que todas as observações deviam contribuir com igual peso para o resultado final. Em qualquer dos casos está presente certa subjetividade na tomada de decisão sobre o que fazer com os *outliers*.  
Fonte: <http://www.ipv.pt/millennium/arq12.htm>.

estatística fornece outra categoria de testes para trabalhar com esses dados, que podem chegar a resultados muito precisos. Esses testes substituem os testes paramétricos, e são denominados como testes não paramétricos. Nesse trabalho, será exemplificado através dos dados da pesquisa TALIS, o teste não paramétrico de Mann-Whitney<sup>20</sup> ou teste U. Vali (2008) informa que esse teste é muito utilizado quando a amostra estudada possui uma grande variação na média, devido à *outliers e sua dispersão (desvio padrão) em volta da média* for muito discrepante (desvio heterogêneo), comprometendo a probabilidade das hipóteses; ou quando os dados forem em escala ordinal, fugindo do padrão de normalidade. Esse teste é considerado um dos testes de maior poder entre os testes não paramétricos e pode ser um teste alternativo ao teste t paramétrico, quando as condições dadas exigirem. Esse teste é indicado, assim como o teste t, para comparação de dois grupos independentes a fim de verificar se pertencem ou não a mesma população. O teste t, testa a igualdade das médias e o teste U testa a igualdade das medianas. Os valores de U calculados pelo teste avaliam o grau de entrelaçamento dos dados dos dois grupos, após a ordenação. A maior separação dos dados em conjunto indica que as amostras são distintas, rejeitando-se a hipótese de igualdade das medianas (VALI, 2008).

A diferença entre os dois testes é que ao invés de construir uma estatística com dados originais, eles são convertidos, para a realização do teste U, em postos ou ordenações. O teste U apresenta como vantagem que as hipóteses de normalidade e homogeneidade da dispersão não são necessárias, permitindo mais generalidade para os resultados. E o *outlier* perde sua influência nesse enfoque, sendo apenas o maior valor da amostra. Outra vantagem, é que ele pode ser aplicado em uma variável que seja originalmente ordinal, enquanto que o teste t exige uma escala pelo menos intervalar. Uma desvantagem é que ao substituir os dados originais por postos, “joga-se fora” alguma informação. Se as condições para o teste t são satisfeitas, ainda assim poder-se-ia usar o teste de Mann-Whitney, mas ele não seria tão poderoso quanto o teste t (VALI 2009).

---

<sup>20</sup>Teste de Mann-Whitney: foi desenvolvido primeiramente por Frank Wilcoxon em 1945, para comparar tendências centrais de duas amostras independentes de tamanhos iguais. Em 1947, Henry Berthold Mann e Emeritus Donald Ransom Whitney generalizaram a técnica para amostras de tamanhos diferentes. Fonte: [http://www.apis2.com.br/?page\\_id=262](http://www.apis2.com.br/?page_id=262)

Existe também a possibilidade de fazer uma comparação múltipla, ou seja, uma técnica estatística que compara várias médias ao mesmo tempo e descobre quais as que diferem entre si (VALI, 2009). O teste F conhecido como ANOVA (análise de variância) é usualmente utilizado para essa finalidade. Esse teste examina a hipótese  $H_0$  na qual três ou mais médias não diferem entre si.

Havendo rejeição dessa hipótese, são realizados testes de comparações múltiplas para saber quais delas diferem entre si. Existem diferentes tipos de testes para comparações múltiplas, como por exemplo, o teste de Tukey (paramétrico). Esse teste consiste em, para cada comparação entre duas médias, comparar a diferença entre elas com a *diferença mínima significativa* (DMS). A regra de decisão é que, se a diferença for maior que a DMS, o teste será significativo e as duas médias consideradas estatisticamente diferentes (OLIVEIRA, 2008).

Segundo OLIVEIRA (2008), a DMS para o teste Tukey é dada por:

$$DMS = q_{(I,v,\alpha\%)} \sqrt{\frac{1}{2} \text{Var}(m_i - m_j)} \quad (10)$$

Em que,

DMS: Diferença mínima significativa;

$q_{(I,v,\alpha\%)}$ : é a amplitude total estudentizada (valor tabelado) para uso no teste Tukey ao nível de  $\alpha\%$  de probabilidade para I tratamentos e v graus de liberdade do Resíduo;

Var: Variância;

$m_i$  e  $m_j$ : tamanho da amostra dos grupos  $i$  e  $j$  respectivamente;

Fonte: Adaptado de OLIVEIRA (2008).

O teste não paramétrico de Kruskal-Wallis<sup>21</sup> (KW) é uma extensão do teste de Mann-Whitney. É um teste não paramétrico utilizado para comparar três ou mais populações independentes. Ele é usado para testar a hipótese  $H_0$  quando todas as medidas de tendência central (média, mediana) não diferem contra a hipótese  $H_1$  de que ao menos uma delas é diferente. Enquanto o teste de Tukey exige que todas as populações em confronto sejam independentes e normalmente distribuídas, o teste de Kruskal-Wallis não impõe nenhuma restrição quanto à distribuição. A fórmula do teste é dada por:

---

<sup>21</sup>Teste criado por William Henry Kruskal (1919 - 2005) e Wilson Allen Wallis (1912-1998).

$$K = (N - 1) \frac{\sum_{i=1}^g n_i (\bar{r}_{i.} - \bar{r})^2}{\sum_{i=1}^g \sum_{j=1}^{n_i} (r_{ij} - \bar{r})^2} \quad (11)$$

Em que,

$K$ : Número de amostra;

$n_i$ : é o número de observações do grupo  $i$ ;

$r_{ij}$ : é a classificação de todas as observações  $j$  no grupo  $i$

$N$  é o número total de observações em todos os grupos;

$$\bar{r}_{i.} = \frac{\sum_{j=1}^{n_i} r_{ij}}{n_i}$$

$\bar{r} = (N + 1)/2$ , é a média de  $r_{ij}$ .

Fonte: adaptado de OLIVEIRA (2008).

Assim, pode-se concluir que: “A inferência estatística é, então um método científico de tirar conclusões sobre os parâmetros da população a partir da recolha, tratamento e análise dos dados de uma amostra, recolhida dessa população” (NEVES, 2009, p. 1).

A estatística inferencial, considerando o erro de estimação, permite dimensionar a confiança que se pode ter no resultado de uma pesquisa ou então subsidiar a tomada de decisão, para evitar a não rejeição ou rejeição de uma hipótese estatística. A partir destes pressupostos, serão realizadas algumas análises e estudos comparativos, a partir de exemplos retirados do relatório, dos questionários e do banco de dados da pesquisa TALIS (OCDE, 2009).

### 3 ANÁLISE DA ESTATÍSTICA NAS PESQUISAS EDUCACIONAIS

“São dados que estas observações são aptas (capazes), que sejam feitas de erros, e que seus erros devam ser corrigidos para nos permitir deter a melhor informação possível a partir dessas observações, nós nos consagramos para aplicar as modificações mínimas e as correções menores possíveis, suficientemente exatos para tirar as observações do impossível e as remeter para o possível” (Tolmaico e Copernicano).

Um dos grandes problemas da estatística aplicada é o fato de pesquisadores imaginarem, ingenuamente, que a estatística é simples e que qualquer um pode administrá-la. Mas, o que se observa em muitos relatórios são indicadores sem objetivos, questionários mal elaborados, dados mal interpretados, conclusões infundadas. A principal causa desses erros é a falta de compreensão dos fundamentos, medidas e técnicas desta ciência.

A estatística oferece instrumentos para avaliar políticas educacionais, que devem ser fundamentados numa concepção de avaliação que, conforme já indicou Gatti (2004), dependem do referencial teórico e da perspectiva epistêmica da pesquisa. Assim, propõe-se aqui analisar o modo de elaboração de questionários, a importância das medidas estatísticas para conhecer a realidade estudada e os cuidados que se devem tomar para fazer sua articulação com as possibilidades de interpretação dos dados, tomando como exemplo a pesquisa TALIS, conforme já anunciado anteriormente.

#### 3.1. ANÁLISE DA ELABORAÇÃO DOS QUESTIONÁRIOS E DAS ESCALAS UTILIZADAS

Os questionários da pesquisa foram enviados aos professores e diretores, segundo a OCDE (2009), de forma impressa ou *online*. Bandeira adverte que:

[um] instrumento autoaplicável (p. ex., enviado via correio, ou *online*), a introdução não somente precisa ser persuasiva, mas deve conter toda a informação necessária para poder agir da maneira esperada pelo pesquisador. Embora se devam incluir indicações claras de como entrar em contato com o responsável pela pesquisa, caso existam dúvidas, o esforço para pedir instruções adicionais pode fazer com que a maioria dos potenciais respondentes ignore o instrumento, em vez de se informar com o pesquisador (BANDEIRA, 2003 p. 4).



Sobre esta consideração, pode-se inferir que existam diferenças de interpretação das questões quando o professor tem a colaboração direta de um aplicador ou precisa resolver sozinho, respondendo por meio de um computador. Em conversa com alguns professores que responderam o questionário da pesquisa TALIS, de duas escolas secundárias da cidade de Jalisco, no México, estes informaram que não sabiam do que se tratava e para qual fim seriam utilizados os dados resultantes do questionário<sup>22</sup>.

Além de o questionário ser muito extenso, causando desconforto e cansaço, demorando em média 45 minutos para responder, as questões possuem muito itens. Das 43 questões dos questionários dos professores, 12 questões possuem um número de itens muito alto, entre 7 e 18 itens. No questionário dos diretores observa-se o mesmo padrão, das 37 questões, 12 possuem um número de itens muito alto, com no mínimo 6 e no máximo 18. Goode e Hatt (1975) advertem que um questionário que a pessoa responde, via correio ou online, não deve exigir mais de 30 minutos e, se possível, um tempo mais curto, porque mesmo com esta duração é difícil obter as respostas desejadas sem cansar o informante. O relatório da OCDE (2009) informa que não havia obrigatoriedade em responder os questionários, mas que o número de perdas de questionários não poderia ser superior a 25%, pois acima desse valor estipulado, poderiam ocorrer problemas com a amostra. Caso esse que ocorreu com a Holanda, que, portanto não pode participar da análise geral apresentada no relatório.

Através das análises dos resultados podem se originar tomadas de decisão, que podem interferir nas políticas educacionais. Porém, toda pesquisa, para atingir seus objetivos, depende de um questionário bem elaborado. Pois, uma pesquisa consiste em uma informação poderosa, no momento em que proporciona reflexões sobre determinado tema, que no caso da TALIS foram importantes para levantar opiniões, percepções e crenças dos professores e diretores, em relação ao processo de ensino e aprendizagem e seu ambiente de trabalho (OCDE 2009). Um bom questionário facilita a organização e interpretação dos dados, por meio da aplicação de técnicas estatísticas. Mas, um questionário mal elaborado impede a aplicação

---

<sup>22</sup>Este contato foi estabelecido em uma missão de estudos da pesquisa O aporte das pesquisas internacionais de avaliação do ensino e da aprendizagem e as políticas públicas de formação e trabalho docente, realizada em setembro de 2013, pelo Grupo de Estudos e Pesquisas em Educação Comparada de Políticas Educacionais (GEPEC), financiada pela Fundação Araucária.

adequada da estatística e pode gerar resultados equivocados. Para evitar problemas ou dúvidas nas análises e cuidados na elaboração do um questionário, Chagas esclarece que:

A construção de um questionário deriva de um processo de melhoria, fruto de tantos exames e revisões quantas forem necessárias. Cada questão deve ser analisada individualmente, para garantir se é mesmo importante, se não é ambígua ou de difícil entendimento, etc. Todas as indagações quanto ao conteúdo, forma, redação e sequência devem ser feitas para cada questão. Uma vez concluída a revisão, feita pela equipe de pesquisa, o questionário estará pronto para o pré-teste. Após revisão originada no pré-teste o questionário estará em condições de ser aplicado eficazmente na pesquisa (CHAGAS, 2000 p.50).

Tendo em vista esta perspectiva, apresentam-se a seguir as possibilidades de análise dos dados da referida pesquisa, através dos questionários aplicados aos professores e diretores, considerando algumas questões sob um olhar estatístico e levantando questionamentos relacionados às escalas de medidas e como estas podem contribuir ou não para um resultado satisfatório. Além disso, alerta-se sobre os cuidados com as conclusões e comparações que podem causar implicações nas políticas públicas de forma equivocada.

A determinação da escala em um questionário tem fundamental importância, pois é através da escala que as análises estatísticas são determinadas. No caso, a questão nº 2 do questionário sobre a idade do professor; assim como a questão nº 9 que pergunta há quanto tempo cada respondente trabalha como professor; e a nº 10, sobre quanto tempo trabalha nesta escola, estão em uma escala ordinal, limitando os recursos da estatística para fazer uma análise mais aprofundada, como segue nos Quadros 04 e 05:

Quadro 4: PERGUNTA DO QUESTIONÁRIO DOS PROFESSORES DA PESQUISA TALIS (2007-08) RELACIONADO À IDADE.

|                                       |                                       |                                       |                                       |                                       |                                       |
|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|
| <b>2. Qual é a sua idade?</b>         |                                       |                                       |                                       |                                       |                                       |
| Menos de 25                           | 25-29                                 | 30-39                                 | 40-49                                 | 50-59                                 | 60+                                   |
| <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> | <input type="checkbox"/> <sub>5</sub> | <input type="checkbox"/> <sub>6</sub> |

Fonte: Questionário dos professores da Pesquisa TALIS, OCDE 2009.

Quadro 5: PERGUNTA DO QUESTIONÁRIO DOS PROFESSORES DA PESQUISA TALIS (2007-08) RELACIONADO AO TEMPO DE TRABALHO DOCENTE.

**9. Há quanto tempo você trabalha como professor?**

*Se possível, exclua períodos prolongados de ausência (como, por exemplo, interrupções na carreira docente)*

Este é meu primeiro ano      1-2 anos      3-5 anos      6-10 anos      11-15 anos      16-20 anos      Há mais de 20 anos

☐<sub>1</sub>      ☐<sub>2</sub>      ☐<sub>3</sub>      ☐<sub>4</sub>      ☐<sub>5</sub>      ☐<sub>6</sub>      ☐<sub>7</sub>

**10. Há quanto tempo você trabalha como professor nesta escola?**

*Se possível, exclua períodos prolongados de ausência (como, por exemplo, interrupções na carreira docente).*

Este é meu primeiro ano      1-2 anos      3-5 anos      6-10 anos      11-15 anos      16-20 anos      Há mais de 20 anos

☐<sub>1</sub>      ☐<sub>2</sub>      ☐<sub>3</sub>      ☐<sub>4</sub>      ☐<sub>5</sub>      ☐<sub>6</sub>      ☐<sub>7</sub>

Fonte: Questionário dos professores da Pesquisa TALIS, OCDE 2009.

As questões citadas nos Quadros 04 e 05, utilizando a escala ordinal, podem ser organizadas conforme as tabelas abaixo. Mas, pode-se apenas fazer descrição das idades em percentuais, conforme Tabela 02.

Tabela 2: PERCENTUAL DE PROFESSORES RELACIONADOS À FAIXA ETÁRIA

| País        | Abaixo de 25 anos | Entre 25 - 29 anos | Entre 30 - 39 anos | Entre 40 - 49 anos | Entre 50 - 59 anos | Acima de 60 anos |
|-------------|-------------------|--------------------|--------------------|--------------------|--------------------|------------------|
|             | %± SE             | %± SE              | %± SE              | %± SE              | %± SE              | %± SE            |
| Brasil      | 6,10 ± 0,80       | 15,90 ± 0,92       | 34,10 ± 1,10       | 31,50 ± 1,00       | 11,20 ± 0,61       | 1,20 ± 0,19      |
| México      | 3,00 ± 0,47       | 11,70 ± 1,01       | 25,80 ± 1,01       | 37,30 ± 1,14       | 18,70 ± 0,94       | 3,50 ± 0,49      |
| Espanha     | 0,40 ± 0,17       | 6,20 ± 4,46        | 29,70 ± 1,06       | 33,80 ± 0,95       | 25,80 ± 1,06       | 4,10 ± 0,45      |
| Média TALIS | 3,00 ± 0,11       | 12,10 ± 0,19       | 28,00 ± 0,23       | 29,60 ± 0,23       | 23,50 ± 0,21       | 3,90 ± 0,10      |

Fonte: OCDE, TALIS 2009, DataBase.

SE: Erro padrão

Ao analisar o percentual de idade dos docentes desse modo, pode-se observar apenas uma tendência de envelhecimento da profissão, que o Brasil, ainda possui uma população docente mais jovem em relação ao México, à Espanha e à Média TALIS, pois 50% dos docentes brasileiros estão entre 25 e 39 anos. Pode-se, ainda, evidenciar um fator preocupante, que é o percentual de docentes com idade

abaixo de 25 anos, 6% para o Brasil, 3% para o México e 0,4% na Espanha, ou seja, que o número de jovens professores está decrescendo (TROJAN; SIPRAKI, 2013).

Mas, não se pode identificar o professor mais novo, nem o mais velho, nem a idade média, nem o grau de dispersão. O conhecimento desses dados permitiria fazer um cruzamento com os dados referentes ao tempo de experiência dos docentes na profissão. Ou seja, não é possível saber se o percentual de professores mais jovens corresponde ao percentual de menos experientes, ou se professores com mais idade estão ingressando na profissão.

Tabela 3: PERCENTUAL DE PROFESSORES RELACIONADOS ÀO REGIME E EXPERIÊNCIA DE TRABALHO (2007-08).

| País        | Experiência de Trabalho |                    |                    |                  |
|-------------|-------------------------|--------------------|--------------------|------------------|
|             | Abaixo de 02 anos       | Entre 03 - 10 anos | Entre 11 - 20 anos | Acima de 21 anos |
|             | %± SE                   | %± SE              | %± SE              | %± SE            |
| México      | 8,70 ± 1,02             | 27,70 ± 1,15       | 29,40 ± 1,27       | 34,20 ± 1,63     |
| Espanha     | 5,80 ± 0,49             | 28,40 ± 1,02       | 30,60 ± 0,91       | 35,20 ± 1,36     |
| Brasil      | 9,60 ± 0,77             | 38,60 ± 1,24       | 32,40 ± 1,09       | 19,30 ± 1,00     |
| Média TALIS | 8,30 ± 0,34             | 29,20 ± 1,13       | 27,00 ± 1,21       | 35,50 ± 1,60     |

Fonte: OCDE, TALIS 2009, DataBase

SE: Erro padrão

Os dados descritivos sobre o percentual médio relacionado à experiência de trabalho na Tabela 03 mostram que no Brasil 49% dos docentes possuem experiência de até 10 anos, enquanto no México e na Espanha esse percentual cai mais de 10%, indicando o primeiro como o país que possui menos professores experientes e que 34,2% no México e 35,2% na Espanha se encontram com mais de 20 anos de trabalho, ou seja, em fase próxima da aposentadoria (TROJAN; SIPRAKI, 2013). Mas, estas análises não permitem estabelecer cruzamento com a idade desses docentes.

Se essas perguntas estivessem em uma escala de razão, poderiam ser realizados levantamentos descritivos importantes. Além de identificar a idade do professor mais novo e o mais velho, ou seja, identificar o limite inferior e o superior e a idade média; analisar as medidas de tendência central como a média, a mediana e

a moda, analisar a dispersão da idade através do desvio padrão e do coeficiente de variação. Também seria possível identificar o professor menos experiente e o mais experiente, fazendo cruzamentos com a idade. E na estatística inferencial, fazer comparações entre as idades dos professores com outros países através dos testes paramétricos de maior poder e confiabilidade, como comparação de intervalos e comparação de médias, já explicados no Capítulo 2.

No questionário de 2012-13 as opções de respostas dessas questões foram alteradas para escala de razão. Assim sendo, ao indagar sobre a idade do professor, pergunta “Qual é a sua idade?” e deixa espaço para o professor colocar a sua idade exata. O mesmo efeito para as questões sobre tempo de experiência na profissão e tempo de trabalho na escola específica. Isso mostra um refinamento da escala de medida e preocupação para uma análise de maior qualidade. Entretanto, não será possível verificar a variação desses dados, comparando os resultados das duas aplicações da pesquisa.

Outro quesito importante a ser discutido é a utilização da escala Likert de 4 pontos, que foi escolhida para identificar as crenças dos professores, sendo 1 para “discordo totalmente”, 2 para “discordo”, 3 para “concordo” e 4 para “concordo firmemente” (OCDE, 2009). A partir das considerações citadas, pode-se inferir que há uma possibilidade de haver indução de determinada tendência. Assim, o estudo de qualquer pesquisa que utilize métodos estatísticos deve analisar os objetivos, a opção teórica e a adequação dos métodos empregados.

Das 43 questões que constam no questionário dos professores, 18 delas estão nessa escala. E para efeito do presente estudo, foram selecionadas duas questões. A questão nº 31, referente às práticas de ensino, crenças e atitudes, e a questão nº 43, sobre a turma específica na qual os professores lecionam, para exemplificar a utilização da escala Likert de 4 pontos, que avalia os níveis de concordância e discordância, isto é, foram oferecidas quatro opções para responder “até que ponto você concorda ou discorda das afirmações a seguir”, salientando que poderia ser marcada apenas uma alternativa em cada linha entre as opções: discordo totalmente, discordo, concordo, concordo totalmente (OCDE, 2009).

A questão nº 31, referente a práticas de ensino, crenças e atitudes, apresenta: .... *sobre você como professor nesta escola*, com seis alternativas e .... *sobre o que acontece nesta escola*, com 4 alternativas como mostra o quadro 05. (OCDE 2009).

Este quadro mostra a questão nº 31, que envolve seis afirmações para avaliar o nível de concordância dos professores sobre: sua satisfação com o trabalho; a sua influência na vida educacional dos alunos; a sua capacidade de motivar os alunos mais difíceis; o sucesso obtido com os alunos; a conquista da confiança dos alunos; o respeito recebido da comunidade, e sobre o que acontece na escola: se os professores e os alunos estão sempre em harmonia; se os professores acreditam que o bem estar dos alunos é importante; se os professores ouvem os alunos; se a escola oferece ajuda extra aos alunos. Para avaliar a escala será analisado a terceira, a quarta e a quinta afirmação referente a questão 31, comparando a descrição dos dados entre Brasil, Espanha, México e a média TALIS.

Quadro 6: PERGUNTA DO QUESTIONÁRIO DOS PROFESSORES DA PESQUISA TALIS (2007-08) REFERENTE ÀS PRÁTICAS DE ENSINO, CRENÇAS E ATITUDES.

| 31. O quanto você concorda ou discorda com as seguintes afirmações ...  |                                       |                                       |                                       |                                       |
|---|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|
| <i>Por favor, marque apenas uma alternativa em cada linha.</i>  |                                       |                                       |                                       |                                       |
| ... sobre você como professor nesta escola?   | Discordo totalmente                   | Discordo                              | Concordo                              | Concordo totalmente                   |
| a) De modo geral, estou satisfeito com meu trabalho. ....   | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| b) Sinto que estou fazendo uma diferença educacional significativa na vida de meus alunos. ..                       | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| c) Se realmente me esforçar, posso conseguir que até meus alunos mais difíceis e desmotivados façam progresso. .... | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| d) Sou bem sucedido com os alunos da minha turma. ....  | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| e) Geralmente, sei como ganhar a confiança dos alunos. ....   | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| f) Os professores nesta comunidade são muito respeitados. ....  | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| ... sobre o que acontece nesta escola?  | Discordo totalmente                   | Discordo                              | Concordo                              | Concordo totalmente                   |
| g) Nesta escola, os professores e os alunos geralmente se dão bem uns com os outros. ....                           | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| h) A maior parte dos professores desta escola acreditam que o bem estar dos alunos é importante. ....               | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| i) A maior parte dos professores desta escola está interessada no que os alunos têm a dizer. ....                   | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| j) A escola oferece ajuda extra se o aluno necessitar. ....   | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |

Fonte: Questionário dos professores da pesquisa TALIS, OCDE 2009.

Na terceira afirmação (letra c), quando perguntados se, mediante esforço pessoal, os professores podem conseguir que até seus alunos mais difíceis façam progresso, a maioria concordou. Isto significa que esta mesma maioria assume que é sua a responsabilidade pelo progresso dos alunos, esquecendo ou minimizando condições importantes que independem de sua vontade, como a condição socioeconômica e cultural dos alunos e a garantia das condições materiais adequadas, que são de responsabilidade do Estado. Causa espanto que os professores concordem que o avanço dos seus alunos depende do seu próprio esforço, ou seja, de sua autoeficácia [sic]. Seria conveniente investigar se as respostas dos professores são condicionadas pelas teorias relacionadas ao construtivismo, às competências e à autoeficácia ou pela formulação da questão e das opções oferecidas.

Tabela 4: AUTOEFICÁCIA: PROGRESSO COM OS ALUNOS (2007-08)

*% de professores do ensino secundário básico, que relataram os seguintes resultados.*

| Países             | Se realmente me esforçar, posso conseguir que até meus alunos mais difíceis e desmotivados façam progresso. |                    |                    |                     |
|--------------------|---|--------------------|--------------------|---------------------|
|                    | Discordo Totalmente   | Discordo           | Concordo           | Concordo Totalmente |
|                    | %± SE   | %± SE              | %± SE              | %± SE               |
| Brasil             | 1,2± 0,22   | 13,7 ± 0,90        | 59,4 ± 0,98        | 25,7 ± 1,07         |
| México             | 0,4 ±0,13   | 4,0 ± 0,60         | 53,2 ± 1,01        | 42,3 ± 1,21         |
| Espanha            | 2,5 ±0,29   | 31,0 ± 1,07        | 54,2 ± 1,09        | 12,2 ± 0,67         |
| <i>TALIS Média</i> | <i>1,3 ±0,02</i>  | <i>18,9 ± 0,11</i> | <i>56,5 ± 0,13</i> | <i>23,4 ± 0,10</i>  |

Fonte: OCDE, Banco de Dados TALIS, 2009.

SE: Erro padrão.

Se observarmos a Tabela 04, verifica-se que os docentes mexicanos (53,2% concordam e 42,3%, concordam totalmente) e brasileiros (59,4% concordam e 25,7% concordam totalmente) são os que apresentam maior percentual de concordância, enquanto os espanhóis (54,2% e 12,2%) se aproximam mais da média TALIS (56,5% e 23,4%). De todo modo, todos os percentuais situam-se acima de 50%. Cabem aqui as considerações de Pointer (2010) ao dizer que a escala Likert permite determinar a direção – positiva, neste caso – do respondente em relação à afirmação, mas, como adverte Garland (1991), a escala de 4 pontos pode

estar reforçando uma tendência e compelindo os respondentes a marcarem a direção mais adequada ou socialmente mais desejável.

Na afirmação seguinte, sobre a percepção do sucesso dos professores em relação aos seus alunos, encontramos a mesma situação de predominância das respostas positivas, ou seja, de concordância em relação à afirmação e, neste caso, os percentuais são ainda mais elevados.

Tabela 5: QUESTÃO 31: AUTOEFICÁCIA: SUCESSO COM OS ALUNOS (2007-08)

% de professores do ensino secundário básico, que relataram os seguintes resultados.

| Países             | Professores relataram que são bem sucedidos com seus alunos |           |           |                        |
|--------------------|---|-----------|-----------|------------------------|
|                    | Discordo<br>Totalmente                                      | Discordo  | Concordo  | Concordo<br>Totalmente |
|                    | %± SE   | %± SE     | %± SE     | %± SE                  |
| Brasil             | 0,3±0,10  | 8,4±0,61  | 73,2±0,98 | 18,1±1,00              |
| México             | 0,2±0,07  | 5,8±0,55  | 72,1±1,01 | 21,9±0,98              |
| Espanha            | 0,4±0,10  | 10,3±0,61 | 80,3±0,78 | 9,1±0,61               |
| <i>TALIS Média</i> | 0,2±0,01  | 7,5±0,06  | 76,1±0,09 | 16,2±0,08              |

Fonte: OCDE, Banco de Dados TALIS.

SE: Erro padrão

Se observarmos a Tabela 05, verificamos que os docentes espanhóis (80,3% e 9,1%) são os que apresentam maior percentual de concordância, enquanto os brasileiros (73,2% concordam e 18,1% concordam totalmente) e os mexicanos (72,1% concordam e 21,9%, concordam totalmente) se aproximam mais da média TALIS (76,1% e 16,2%).

Contudo, mais que analisar os resultados, cabe inquirir o que pode significar para os professores *ser bem sucedido com os alunos da sua turma*. O professor respondeu que conseguiu êxito no processo de ensino e aprendizagem ou que possui popularidade com os alunos? Além do mais, até que ponto um professor admitiria não ser bem sucedido, qualquer que seja a interpretação de sucesso? Se considerarmos a opção referente ao êxito na docência, seria conveniente verificar o desempenho escolar dos estudantes. De todo modo, não será uma avaliação adequada se não forem ponderadas as condições socioeconômicas dos estudantes e os recursos materiais e pedagógicos disponíveis para a realização do processo educativo.



Em outra afirmação, sobre o quanto os professores estão fazendo diferença na vida educacional dos alunos, a análise feita no relatório apresenta que a maioria (mais de 90%) acredita que faz diferença, como mostra na Tabela 06 (OCDE, 2009).

Tabela 6: QUESTÃO 31: AUTOEFICÁCIA: FAZENDO A DIFERENÇA SIGNIFICATIVA EDUCACIONAL DOS ALUNOS (2007-08).

*% de professores do ensino secundário básico, que relataram os seguintes resultados.*

| Países             | Professores relataram fazer diferença significativa na vida de seus alunos |           |           |                     |
|--------------------|--|-----------|-----------|---------------------|
|                    | Discordo Totalmente  | Discordo  | Concordo  | Concordo Totalmente |
|                    | %± SE  | %± SE     | %± SE     | %± SE               |
| Brasil             | 0,8±0,18   | 7,8±0,67  | 71,1±1,08 | 20,3±1,11           |
| México             | 0,3±0,15   | 2,5±0,38  | 61,2±1,16 | 36,0±1,25           |
| Espanha            | 0,5±0,12   | 10,4±0,55 | 75,6±0,81 | 13,5±0,71           |
| <i>TALIS média</i> | 0,5±0,02   | 8,9±0,06  | 64,1±0,11 | 26,4±0,10           |

Fonte: OECD, Banco de Dados TALIS 2009.

SE: Erro padrão

Mas, para chegar a essa conclusão, foram somadas duas respostas (concordo 64,1% e concordo totalmente 26,4% na média TALIS). Há de se pensar que apenas 26,4% concordam plenamente. Para evitar que sejam prejudicadas as comparações entre os países e a avaliação de políticas educacionais, que utilizem os dados da pesquisa, seria prudente replicar estas questões, incluindo uma alternativa neutra para avaliar a força dos resultados obtidos.

Outro ponto a considerar é que as afirmativas colocadas para avaliar as práticas de ensino, crenças e atitudes dos professores desconsideram os aspectos relacionados às políticas de financiamento da educação e de formação docente e à organização e gestão dos sistemas de ensino. Ao que parece, a intenção é atribuir somente ao professor a responsabilidade sobre os resultados do processo de ensino e aprendizagem.

A questão nº 43, referente a turma específica, está representada como: *o quanto você discorda ou concorda com cada uma das afirmações a seguir no que diz respeito a turma específica*, com 4 alternativas para responderem marcando apenas uma alternativa em cada item como mostra o (Quadro 07). (OCDE 2009).

Quadro 7: PERGUNTA DO QUESTIONÁRIO DOS PROFESSORES DA PESQUISA TALIS (2007-08). REFERENTE À TURMA ESPECÍFICA.

**43 O quanto você discorda ou concorda com cada uma das afirmações a seguir no que diz respeito a essa turma específica?**

*Por favor, marque apenas uma alternativa em cada linha.*

|  | Discordo totalmente                   | Discordo                              | Concordo                              | Concordo totalmente                   |
|--|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|
| a) No início da aula, tenho que esperar muito tempo até que os alunos se acalmem. ....   | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| b) Os alunos dessa turma ajudam a criar uma atmosfera agradável para o aprendizado. .... | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| c) Eu perco muito tempo porque os alunos interrompem a aula. ....                        | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| d) Há muito barulho nessa sala de aula. ....   | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |

Fonte: Questionário dos professores da pesquisa TALIS, OCDE 2009.

No que se refere ao ensino para uma turma específica, no Quadro 07 são apresentadas quatro afirmações para avaliar a concordância ou discordância do professor em relação aos problemas que podem prejudicar o processo de ensino e aprendizagem, utilizando a escala de quatro pontos, com as mesmas opções das anteriores. Nesse caso, pergunta o quanto o professor discorda ou concorda com as seguintes afirmações: a) No início da aula, tenho que esperar muito tempo até que os alunos se acalmem; b) Os alunos dessa turma ajudam a criar uma atmosfera agradável para o aprendizado; c) Eu perco muito tempo porque os alunos interrompem a aula; e, d) Há muito barulho nessa sala de aula (OCDE, 2009). Como mostra a Tabela 7.

Tabela 7: QUESTÃO 43: CLIMA DISCIPLINAR EM SALA DE AULA: BARULHO EM SALA DE AULA (2007-08)

% de professores do ensino secundário básico, que relataram os seguintes resultados.

| Países      | Há muito barulho na sala de aula |             |             |                     |
|-------------|----------------------------------|-------------|-------------|---------------------|
|             | Discordo Totalmente              | Discordo    | Concordo    | Concordo Totalmente |
|             | %± SE                            | %± SE       | %± SE       | %± SE               |
| Brasil      | 8,9 ± 0,66                       | 52,2 ± 1,29 | 32,0 ± 1,08 | 6,8 ± 0,48          |
| México      | 26,9 ± 1,17                      | 55,8 ± 1,13 | 14,0 ± 0,90 | 3,3 ± 0,35          |
| Espanha     | 15,0 ± 0,84                      | 46,6 ± 1,05 | 27,6 ± 1,01 | 10,8 ± 0,60         |
| TALIS Média | 25,2 ± 0,11                      | 48,4 ± 0,12 | 21,0 ± 0,11 | 5,4 ± 0,05          |

Fonte: OECD, Banco de Dados TALIS 2009

SE: Erro padrão

Tomando como exemplo a última afirmação, verifica-se que a maioria dos professores investigados discorda ou discorda totalmente que “Há muito barulho em sala de aula”, como se pode observar na Tabela 07.

Ao analisar os resultados das opções *discordo totalmente* e *discordo*, observa-se que, nos três países selecionados – Brasil (8,9 e 52,2%), México (26,9 e 55,8%) e Espanha (15 e 46,6%), respectivamente –, a grande maioria discorda que ocorra muito barulho em sua sala de aula, resultado semelhante à média TALIS (25,2 e 48,4%).

Entretanto, se confrontarmos essa questão, com a análoga do questionário do diretor, encontra-se resultados totalmente opostos. Ou seja, a maioria dos diretores afirma que o ensino está sendo prejudicado por *perturbação em sala de aula*, como mostra a Tabela 08.

Tabela 8: QUESTÃO 30: EM QUE MEDIDA O ENSINO ESTÁ SENDO PREJUDICADO

| Países      | Perturbação em sala de aula |      |
|-------------|-----------------------------|------|
|             | Média Percentual            | SE   |
| México      | 72,00                       | 3,60 |
| Espanha     | 70,00                       | 4,00 |
| Brasil      | 60,00                       | 3,00 |
| Média TALIS | 60,20                       | 0,80 |

Fonte: OCDE, Relatório TALIS, 2009.

Para Bandeira (2003 p.7) a escolha da palavra pode direcionar uma resposta. E esta cita um exemplo, quando se pergunta a um entrevistado sobre utilização de áreas comuns num bloco de residência, pode-se indagar se deve ser ‘proibido’, ‘não permitido’, ‘evitado’ ou ‘impedido’. Assim como o aviso ‘Proibido Estacionar’ ou ‘Pede-se não estacionar’ provoca comportamentos diferentes, e pode variar na resposta.

Isso pode ter acontecido nos questionários quando perguntado, no item d da questão nº 43, no questionário dos professores e, no item c da questão nº 30, no questionário dos diretores, como mostra o Quadro 08:

Quadro 8: PERGUNTA DO QUESTIONÁRIO DOS PROFESSORES DA PESQUISA TALIS (2007-08), REFERENTE À TURMA ESPECÍFICA AOS RECURSOS DA ESCOLA.

**30. Nesta escola, em que medida o ensino dos alunos está sendo prejudicado pelos seguintes comportamentos?**

*Por favor, marque apenas uma alternativa em cada linha.*

| <b>Por parte dos alunos da escola:</b>   | De forma alguma                       | Muito pouco                           | Até certo ponto                       | Muito                                 |
|--|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|
| a) Chegar atrasado à escola .....  | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| b) Absenteísmo (por exemplo, faltas injustificadas) .....                                | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| c) Perturbação em sala de aula .....   | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| d) Colar .....   | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| e) Profanação/Palavrões .....  | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| f) Vandalismo .....  | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| g) Furto .....   | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| h) Intimidação ou ofensa verbal de outros alunos (ou outras formas de intimidação) ..... | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| i) Danos físicos a outros alunos .....   | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| j) Intimidação ou ofensa verbal a professores ou membros da equipe escolar .....         | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |
| k) Uso/posse de drogas e/ou bebidas alcoólicas .....                                     | <input type="checkbox"/> <sub>1</sub> | <input type="checkbox"/> <sub>2</sub> | <input type="checkbox"/> <sub>3</sub> | <input type="checkbox"/> <sub>4</sub> |

Fonte: Questionário dos diretores da pesquisa TALIS, OCDE 2009.

A caracterização das opções de resposta “de forma alguma; muito pouco; até certo ponto; e muito” pode apresentar entendimento diferente de “discordo totalmente, discordo, concordo e concordo totalmente”. E em termos de escala, permite estabelecer, gradativamente, a medida do barulho, que também pode ser interpretada como um escala ordinal.

Ao confrontar as duas questões do questionário do professor e do questionário do diretor percebe-se que há uma diferença de interpretação pelos respondentes. Vale ressaltar que, essa mesma questão está presente no questionário da escola, aplicado pelo *Programa Internacional de Avaliação de Estudantes* (PISA) em 2009 (OCDE, 2010). Assim, pergunta-se: Por que não foram apresentadas as mesmas opções para os professores?

As opções oferecidas ao professor não apresentam a mesma possibilidade de medida, porque não há como estabelecer qual é a intensidade das opções concordo e discordo. Se: *concordo totalmente* e *discordo totalmente*, correspondem a 100% de peso, qual é o grau que se pode estabelecer para as duas outras opções? Assim

sendo, *concordo* significa menos de 100% de concordância. Mas, quanto? E, ao contrário, qual é o percentual que se pode atribuir à opção *discordo*? Se as opções *concordo* e *discordo* corresponderem a 50%, elas se anulam, significam a mesma coisa. Nesse caso, quando se somam os resultados de *concordo* totalmente e *concordo* ou *discordo* totalmente e *discordo* ignora-se a diferença de intensidade entre as escalas. Esta conclusão também corrobora com as conclusões dos estudos de Garland (1991).

Conforme observado, no relatório TALIS (OCDE, 2009), as tabelas, referentes aos fatores relacionados aos alunos que prejudicam o ensino, apresentam os resultados somados em dois grupos, *discordo* totalmente e *discordo* como uma única alternativa e o mesmo para as opções *concordo* e *concordo* totalmente. Assim, porque utilizar a Escala Likert de quatro pontos se a Escala de Thurstone (OLIVEIRA, 2001) propõe exatamente a escolha entre duas alternativas – *concordo* e *discordo*? Qual é a intenção? Se a proposição de quatro ou mais opções pressupõe uma diferença significativa de intensidade entre as alternativas, estatisticamente, não se podem apresentar somatórias resultantes da junção de respostas, pois, cada uma representa um grau de valor.

Retomando o problema da *perturbação em sala de aula*, no caso do questionário do diretor, o relatório também apresenta os resultados somados, apesar das opções indicarem um grau de intensidade distinto e objetivo, além de não permitir identificar quais foram as opções somadas para afirmar que, na maioria dos países pesquisados, os distúrbios em sala prejudicam o ensino (OCDE, 2009).

Outra consideração importante é a contradição apresentada pelas respostas dos professores e dos diretores. Enquanto a maioria dos diretores afirma esse problema – Brasil (60,2%), Espanha (70,5%), México (71,9%) e Média TALIS (60,2%) –, a maioria dos professores discorda que há muito barulho em sala de aula – Brasil (8,9 e 52,2%), México (26,9 e 55,8%), Espanha (15 e 46,6%) e média TALIS (25,2 e 48,4%), respectivamente em relação a *discordo* totalmente e *discordo*. É possível que a diferença esteja relacionada com a forma de apresentação da questão e das opções de escolha. Outra possibilidade é colocar em dúvida o entendimento da pergunta pelo professor ou o grau de subjetividade que esta encerra.

Assim, se observa da outras afirmações referentes à mesma questão verifica-se que há resultado similar de discordância em relação às afirmações “No início da aula, tenho que esperar muito tempo até que os alunos se acalmem” e “Eu perco muito tempo porque os alunos interrompem a aula” se comparado à questão do barulho. Também não se deve desconsiderar a interpretação do professor sobre o conceito de “muito barulho”.

Tabela 9: COMPARAÇÃO ENTRE VARIÁVEIS DA QUESTÃO Nº 30 - QUESTIONÁRIO DO PROFESSOR.

| Países      | Muito barulho em sala |          | Aguardando alunos se acalmarem |          | Interrupções em sala |          |
|-------------|-----------------------|----------|--------------------------------|----------|----------------------|----------|
|             | Discordo totalmente   | Discordo | Discordo totalmente            | Discordo | Discordo totalmente  | Discordo |
| Brasil      | 8,9                   | 52,2     | 11,3                           | 47,2     | 8,9                  | 56,6     |
| México      | 26,9                  | 55,8     | 36,8                           | 47,6     | 23,5                 | 59,0     |
| Espanha     | 15,0                  | 46,6     | 15,7                           | 43,0     | 12,1                 | 44,9     |
| TALIS Média | 25,2                  | 48,4     | 21,5                           | 42,4     | 17,0                 | 52,9     |

Fonte: OECD, TALIS 2009.

A partir das considerações citadas, pode-se deduzir que há possibilidade de haver indução de determinada tendência. Assim, o estudo de qualquer pesquisa que utilize métodos estatísticos de escala Likert deve analisar os objetivos, a opção teórica e a adequação dos métodos empregados.

Para essas questões formadas através da escala Likert, e confrontadas as questões nº 43 do questionário dos professores no item d “há muito barulho em sala de aula”, com a questão nº 30 do questionário do diretor no item c com a mesma pergunta fica evidente que houve má interpretação da escala Likert. Importante analisar essa questão e reformular a resposta dos itens de “discordo totalmente, discordo, concordo e concordo totalmente” para “de forma alguma; muito pouco; até certo ponto; muito”. Outro problema da escala Likert de 4 pontos obriga o entrevistado optar por uma questão, para esse tipo de questionário poderia ter sido usada a escala Likert de 5 ponto com uma coluna de indiferente.

Com referência a esta questão, não houve alteração no questionário aplicado em 2012-13 aos professores, ou seja, mantém-se os mesmos problemas que

impedem uma interpretação adequada. Mas, estranhamente, a questão correlata do questionário dos diretores – que, segundo a análise aqui realizada, com base em Garland (1991), que estava bem elaborada – foi excluída.

### 3.2. ANÁLISE SOBRE USO DO DESVIO E DO ERRO PADRÃO

Analisando as tabelas descritivas apresentadas no relatório da OCDE 2009, verifica-se que há problemas que precisam ser destacados, no que se refere à indicação do desvio padrão.

Tabela 10: NÚMERO DE ALUNOS MATRICULADOS POR ESCOLA DO BANCO DE DADOS OCDE 2009.

| País    | Média | Desvio Padrão | Mediana | Limite Mínimo | Limite Máximo | Coef. Variação |
|---------|-------|---------------|---------|---------------|---------------|----------------|
| Brasil  | 880   | 636           | 775     | 14            | 4141          | 72             |
| Espanha | 668   | 391           | 600     | 45            | 2118          | 59             |
| México  | 554   | 423           | 429     | 17            | 1898          | 76             |

Fonte: OCDE, Banco de Dados TALIS, 2009.

De acordo com o relatório OCDE 2009, não é possível identificar o porte das escolas que fizeram parte da pesquisa. Porém, a partir do questionário dos diretores, na questão 12, pode-se identificar o tamanho da escola a partir do número de alunos matriculados, como indica a Tabela 10. Se for analisada somente a média – como é de costume na análise das pesquisas educacionais –, nota-se que o Brasil possui um número médio de 880 alunos matriculados por escola, assim como a Espanha possui 668 e o México 554. Nesta perspectiva, pode-se levantar a hipótese de que as escolas que participaram da pesquisa possuem um porte relativamente alto. Mas, quando confrontado com o desvio padrão, observa-se que existe uma dispersão extremamente alta em torno da média, isto é, existe uma grande heterogeneidade entre o tamanho das escolas na pesquisa. Os limites mínimos do Brasil, da Espanha e do México são de 14, 45 e 17 alunos matriculados respectivamente, ou seja, escolas extremamente pequenas, e um limite máximo 4.141, 2.118 e 1.898 alunos

matriculados nas escolas de grande porte. A dispersão do desvio padrão pode ser notada através do coeficiente de variação, que indica o percentual de dispersão em torno da média de 72% para o Brasil, 59% para a Espanha e 76% para o México, o que evidencia o equívoco na utilização da média. Neste caso, deveria ter sido utilizada a mediana que apresenta uma diferença de quase 100 alunos em relação a média, nos casos do Brasil e do México e 68 alunos, no caso da Espanha.

No caso do número de alunos por turma, a Tabela 11 do relatório TALIS não apresenta o desvio padrão necessário para identificar se existe dispersão significativa em relação à média, inclui somente o erro padrão.

Tabela 11: MÉDIA DE ALUNOS POR TURMA DO RELATÓRIO DA OCDE 2009.

| Países      | Número médio de alunos por turma |                  |
|-------------|----------------------------------|------------------|
|             | Média                            | Erro padrão (SE) |
| México      | 38                               | 0,55             |
| Espanha     | 22                               | 0,26             |
| Brasil      | 35                               | 0,35             |
| Média TALIS | 24                               | 0,07             |

Fonte: OCDE, Relatório TALIS, 2009.

O erro padrão é uma medida a qual reflete a precisão da estimativa obtida para a média de alunos por turma. Esta precisão aumenta conforme o tamanho da amostra cresce. Não se trata de uma medida de variabilidade entre as turmas de alunos (como o desvio padrão), mas sim de uma medida de precisão da estimativa da média, em função do tamanho da amostra. Assim, quanto maior for a amostra, menor será o erro padrão, pois:

No caso do erro padrão da média, este é obtido dividindo o desvio padrão da amostra pela raiz quadrada do número de observações na amostra.

O erro padrão da estimativa diminui com o aumento do tamanho da amostra, refletindo o aumento de precisão da estimativa com o tamanho da amostra (LUNET; SEVERO; BARROS, 2006, p. 55).

Entretanto, o levantamento de dados, feito diretamente do banco fornecido pelo site da OCDE, permite identificar a média (Mean), o tamanho da amostra (N), o desvio padrão amostral (Std. Deviation), a mediana (Median), o erro padrão da média (Std. Error of Mean), o limite inferior (Minimum), o limite superior (Maximum), e



o coeficiente de variação, tal como descreve a (Tabela 12), desenvolvida por meio do programa estatístico SPSS 13.0<sup>23</sup>.

Tabela 12: DESCRIÇÃO DO NÚMERO DE ALUNOS POR TURMA DO BANCO DE DADOS OCDE.

| Alunos por turma | Mean  | N     | Std. Deviation | Median | Std. Error of Mean | Minimum | Maximum | Coef. Var. |
|------------------|-------|-------|----------------|--------|--------------------|---------|---------|------------|
| Brasil           | 34,00 | 5491  | 10,7           | 35,00  | 0,145              | 1       | 98      | 31%        |
| Espanha          | 21,8  | 3263  | 7,7            | 23,00  | 0,135              | 1       | 95      | 35%        |
| México           | 39,2  | 3101  | 15,6           | 38,00  | 0,28               | 4       | 98      | 40%        |
| Itália           | 21,5  | 4936  | 4,88           | 22,00  | 0,07               | 1       | 98      | 23%        |
| Total            | 28,8  | 16791 | 12,41          | 26,00  | 0,09               | 1       | 98      | 43%        |

Fonte: Base de dados da Pesquisa TALIS, OCDE 2009

Nota-se que os desvios padrão referente aos países Brasil e México se apresentam relativamente altos em relação à média amostral, ou seja, a dispersão entre as escolas em torno da média é heterogênea. Porém, como os tamanhos amostrais são distintos, assim como as médias, seria conveniente usar o coeficiente de variação para comparar essa dispersão em torno da média. Através do coeficiente de variação pode-se verificar que o México (40%) foi o país que apresentou maior dispersão em torno da média, seguido da Espanha (35%), do Brasil (31%) e da Itália (23%). De acordo com a classificação de Mohallen, (2008), explicado no Capítulo 2, existe uma alta dispersão em torno da média, exceto para a Itália, cuja dispersão pode ser classificada como moderada.

Todavia, todas as tabelas apresentadas no relatório da OCDE mostram apenas o erro padrão da média (*Std. Error of Mean*), e não o desvio padrão (*Std. Deviation* ou *SD*). Nota-se que o desvio padrão seria mais relevante para o leitor leigo em estatística (que, em geral, desconhece a relação entre desvio padrão e erro padrão) para identificar a dispersão em torno da média. A indicação do erro padrão é importante para avaliar a amostra. Assim,

<sup>23</sup>O chamado programa SPSS (*Statistical Package for Social Sciences*) é uma ferramenta para análise de dados, que permite utilizar técnicas estatísticas básicas e avançadas. Segundo Mundstocket al (2006): “É um *software* estatístico de fácil manuseio internacionalmente utilizado há muitas décadas, desde suas versões para computadores de grande porte”. Fonte: [http://www.mat.ufrgs.br/~camey/SPSS/Introdu%20o%20An%20lise%20Estat%20stica%20utilizando%20o%20SPSS%2013\\_0.pdf](http://www.mat.ufrgs.br/~camey/SPSS/Introdu%20o%20An%20lise%20Estat%20stica%20utilizando%20o%20SPSS%2013_0.pdf)

Se o objetivo é descrever a variabilidade observada numa amostra deve-se utilizar o desvio padrão. O desvio padrão, como medida de dispersão, não deve ser usado quando a população não segue uma distribuição normal ou aproximadamente normal. Nestes casos, o desvio padrão pode não ser uma boa estimativa de dispersão, pelo fato da média, que é utilizada no seu cálculo, ser pouco resistente a observações extremas (LUNET; SEVERO; BARROS, 2006, p. 56).

Nesse caso, a ausência do desvio padrão e coeficiente de variação nas tabelas do relatório da OCDE, oculta a alta dispersão dos dados em torno da média. O erro padrão tem outro objetivo, qual seja:

[...] indicar a imprecisão associada à estimativa de um determinado parâmetro (e.g. média), pode utilizar-se o erro padrão. Contudo, de uma forma geral, os intervalos de confiança podem ser interpretados de forma mais direta que os erros padrão, sendo preferível a apresentação dos primeiros. O erro padrão é um passo intermédio no cálculo de intervalos de confiança. O fato do erro padrão ser quantitativamente menor do que o desvio padrão pode contribuir para que alguns autores optem por apresentar o erro padrão quando pretendem quantificar a dispersão das observações da amostra, transmitindo uma falsa ideia de precisão aos leitores menos atentos e com poucos conhecimentos de estatística. (LUNET; SEVERO; BARROS, 2006, p. 56).

Assim, é fundamental escolher a medida de variabilidade mais informativa para apresentar aos leitores nos relatórios, o que não acontece no relatório da OCDE. Tendo em vista que se trata de uma pesquisa em larga escala, o que implica em tamanhos amostrais grandes, é esperado que o erro padrão seja pequeno, o que não ocorre com o desvio padrão que informa a variabilidade dos dados em torno da média. De todo modo, o erro padrão que consta nas tabelas não é utilizado para análise no relatório, permitindo equívocos na interpretação dos dados.

[...] se é apresentada a média e a respectiva medida de incerteza, os resultados são a média e o erro padrão, ou se a média é uma medida de dispersão, os resultados apresentados são a média e o desvio padrão. (LUNET; SEVERO; BARROS, 2006, p. 56).

Também se verifica que o cálculo do erro padrão nas Tabelas 10 (do banco de dados da OCDE) e 11 (do relatório da OCDE) não apresentam os mesmos resultados<sup>24</sup>. Finalmente, comparando essas mesmas tabelas verifica-se que a tabela 11 – que apresenta somente o erro padrão – indica que a média está sendo estimada com alta precisão devido ao tamanho amostral grande, e a tabela 10 – que

---

<sup>24</sup>Nesse caso, não se pode afirmar o que ocorreu. Não é possível identificar se foi feito algum tipo de exclusão, como por exemplo, os *outliers*.

utiliza o desvio padrão amostral – revela que há uma dispersão expressiva entre as médias das escolas de cada país em relação a sua média geral, que é ignorada no relatório, permitindo equívocos de interpretação. Por esta razão, é fundamental compreender a diferença entre erro padrão e desvio padrão.

Além das análises descritivas apresentadas, pesquisas internacionais como a TALIS possuem a intenção de permitir estudos comparados sobre a condição de um país em relação a outro, em relação a cada variável investigada. Assim, propõe-se desenvolver algumas comparações a partir de medidas de tendência central, com a finalidade de subsidiar o uso da estatística nas pesquisas educacionais.

### 3.3. COMPARAÇÃO DE MEDIDAS DE TENDÊNCIA CENTRAL

Em muitas pesquisas é comum que se tenha interesse na comparação de medidas de tendência central. Na pesquisa TALIS, por exemplo, pode-se desejar comparar se a média de alunos por turma obtida por dois países (Brasil e Espanha, ou Brasil e México, por exemplo) diferem estatisticamente entre si.

Como os valores obtidos para as médias são baseadas em amostras extraídas das populações de interesse, se faz necessário o uso de testes de comparação ou obtenção de intervalos de confiança a fim de, com base nas amostras observadas, permitir a realização de inferências a respeito de possíveis diferenças estatísticas entre essas medidas, para duas ou mais populações comparadas.

#### 3.3.1. Comparação com base em testes paramétricos entre duas amostras independentes

Para ilustração do teste  $t$  de Student (paramétrico), foram utilizadas duas amostras independentes extraídas do banco de dados da pesquisa TALIS, em três exemplos: comparação entre Brasil e Espanha; Brasil e México; e Espanha e Itália<sup>25</sup>.

Para comparação de médias de alunos por turma entre Brasil e Espanha. As hipóteses testadas foram:  $H_0$ : A média de alunos por turma do Brasil *não difere* da

---

<sup>25</sup> Neste exemplo, foi escolhida a Itália por este país apresentar resultados parecidos com os da Espanha, neste quesito, o que não ocorre nos casos do Brasil e do México.

média de alunos por turma da Espanha; e  $H_1$ : A média de alunos por turma do Brasil *difere* da média de alunos por turma da Espanha.

Tabela 13: DESCRIÇÃO DOS DADOS RELACIONADOS À COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE BRASIL E ESPANHA PARA O TESTE PARAMÉTRICO.

| Group Statistics |         |      |       |                |                 |
|------------------|---------|------|-------|----------------|-----------------|
|                  | país    | N    | Mean  | Std. Deviation | Std. Error Mean |
| Alunos por turma | Brasil  | 5491 | 33,82 | 10,712         | ,145            |
|                  | Espanha | 3263 | 21,75 | 7,696          | ,135            |

Fonte: Base de dados da pesquisa TALIS, OCDE 2009.

A Tabela 13 apresenta as medidas descritivas dos dados, identificando a amostra do Brasil de 5.491 professores, com média de 34 alunos por turma e desvio padrão amostral de 11 alunos por turma, o erro padrão da média é de 0,15. A Espanha possui uma amostra de 3.263 professores, com uma média de 22 alunos por turma, e desvio padrão amostral de 8 alunos por turma, com um erro padrão da média de 0,14. O desvio padrão de 11 alunos do Brasil mostra que há uma heterogeneidade maior de alunos por turma, quando comparado com 8 da Espanha.

Esses dados apenas descrevem a amostra, mas não explicam se as médias entre os dois países são iguais ou diferentes. Não se pode dizer com segurança se o Brasil possui um número maior de alunos por turma que a Espanha, devido a fatores relacionados à dispersão (desvio padrão alto em relação à média) e ao tamanho da amostra dos dois países.

Para obter uma conclusão precisa, necessita-se do teste de comparação das médias para testar as hipóteses e analisar o valor  $p$ , conforme a Tabela 14 sobre a comparação de médias calculadas pelo teste paramétrico  $t$  de Student e a comparação de médias por intervalos.

Primeiramente, o programa SPSS 13.0 faz uma análise das variâncias<sup>26</sup> amostrais (dispersão), apresentando as saídas possíveis para identificar o tipo de teste  $t$  que deve ser utilizado. No exemplo citado, para esse teste, o programa

<sup>26</sup> Variância: é o desvio quadrático médio da média, ou seja, é o desvio padrão ao quadrado. Nas análises estatísticas inferenciais utilizam-se as variâncias. Fonte: <http://leg.ufpr.br/~silvia/CE003/node16.html>.

utilizou o *Teste de Levene*<sup>27</sup>, ou *Teste F*, para comparar a homogeneidade das variâncias, ou seja, se as variâncias diferem ou não entre si. O resultado do teste foi  $F = 99,369$  com valor  $p \approx 0$ , o que significa que não é possível concluir que as variâncias sejam homogêneas. Para isso, o teste deveria apresentar um valor  $p$  superior a 0,05 (grau de significância), conforme explicado anteriormente. Como mostra a Tabela 14.

Tabela 14: ANÁLISE DA COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE BRASIL E ESPANHA, PARA O TESTE PARAMÉTRICO.

| Independent Samples Test |                             |   |      |                              |          |                 |                 |                       |   |        |
|--------------------------|-----------------------------|---|------|------------------------------|----------|-----------------|-----------------|-----------------------|---|--------|
|                          |                             | Levene's Test for Equality of Variances |      | t-test for Equality of Means |          |                 |                 |                       |   |        |
|                          |                             | F                                       | Sig. | t                            | df       | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference |        |
| alunos                   | Equal variances assumed     | 99,369                                  | ,000 | 56,324                       | 8752     | ,000            | 12,074          | ,214                  | 11,654                                    | 12,494 |
|                          | Equal variances not assumed |   |      | 61,101                       | 8445,376 | ,000            | 12,074          | ,198                  | 11,687                                    | 12,461 |

Fonte: Base de dados da Pesquisa TALIS, OCDE 2009.

Logo, o teste  $t$  a ser considerado é o que assume variâncias não homogêneas, ou seja, variâncias diferentes. Sendo assim, o resultado do teste  $t$  que será analisado é o da 2ª linha na Tabela 15 – das saídas (*Equal variances not assumed*). O resultado foi  $t = 8445,376$ , valor  $p \approx 0$  e mostra que é possível rejeitar a hipótese  $H_0$  (as médias não diferem) a favor da hipótese alternativa  $H_1$  (as médias diferem).

Em relação ao intervalo de confiança fornecido, como este não inclui o valor zero, é possível rejeitar  $H_0$  (as médias não diferem) e concluir que as médias diferem ao nível de significância de 5% (ou equivalentemente, um nível de confiança de 95%).

Além disso, o teste informa que a diferença entre as médias desses dois países (Brasil e Espanha) foi de 12,074 alunos, podendo variar com 95% de confiança entre 11,687 e 12,461 alunos. Ou seja, o Brasil tem, em média, 12 alunos a mais por turma do que a Espanha.

<sup>27</sup> Elaborado por Levene em 1960, este teste consiste em fazer uma transformação dos dados originais e aplicar aos dados transformados. Fonte: <http://www.portalection.com.br/content/161-teste-de-igualdade-das-vari%C3%A2ncias>

A vantagem de apresentar o teste t é que com ele pode-se dizer, com base no valor p, até que nível de significância é possível rejeitar  $H_0$ . Aqui o teste t forneceu valor próximo de zero, logo é possível concluir que a diferença entre as médias vale para qualquer nível de significância maior que o valor p encontrado (1%, 5%, 10%, etc).

No segundo exemplo, as hipóteses testadas para comparação de médias entre Brasil e México foram:  $H_0$ : A média de alunos por turma do Brasil *não difere* da média de alunos por turma do México; e  $H_1$ : A média de alunos por turma do Brasil *difere* da média de alunos por turma do México.

Tabela 15: DESCRIÇÃO DOS DADOS RELACIONADOS À COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE BRASIL E MÉXICO, PARA O TESTE PARAMÉTRICO.

| Group Statistics |        |      |       |                |                 |
|------------------|--------|------|-------|----------------|-----------------|
|                  | país   | N    | Mean  | Std. Deviation | Std. Error Mean |
| Alunos por turma | Brasil | 5491 | 33,82 | 10,712         | ,145            |
|                  | México | 3101 | 39,17 | 15,567         | ,280            |

Fonte: Base de dados da pesquisa TALIS, OCDE 2009.

A Tabela 15 apresenta as medidas descritivas dos dados identificando uma amostra para o Brasil de 5.491 professores, com média de 34 alunos por turma e desvio padrão de 11 alunos por turma, o erro padrão amostral é de 0,15. O México possui uma amostra de 3.101 professores, com uma média de 39 alunos por turma, e seu desvio padrão de 16 alunos por turma, com um erro padrão amostral de 0,28.

Tabela 16: ANÁLISE DA COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE BRASIL E MÉXICO PARA O TESTE PARAMÉTRICO.

| Independent Samples Test |                             |   |      |                              |          |                 |                 |                       |  |
|--------------------------|-----------------------------|---|------|------------------------------|----------|-----------------|-----------------|-----------------------|--|
|                          |                             | Levene's Test for Equality of Variances |      | t-test for Equality of Means |          |                 |                 |                       |  |
|                          |                             | F                                       | Sig. | t                            | df       | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference<br>Lower Upper |
| alunos                   | Equal variances assumed     | 267,043                                 | ,000 | -18,765                      | 8590     | ,000            | -5,345          | ,285                  | -5,903 -4,787  |
|                          | Equal variances not assumed |   |      | -16,984                      | 4786,448 | ,000            | -5,345          | ,315                  | -5,962 -4,728  |

Fonte: Base de dados da Pesquisa TALIS, OCDE 2009.

A Tabela 16 diz respeito à comparação de médias calculadas pelo teste paramétrico t de Student e a comparação de médias por intervalos. O teste de Levene, para comparar se as variâncias não diferem, foi de  $F = 267,043$  com valor  $p \approx 0$ , o que significa que não é possível concluir que as variâncias sejam homogêneas, ou seja, rejeita-se a hipótese das variâncias serem iguais ( $H_0$ ). Sendo assim, o resultado do teste t que será utilizado é o da 2ª linha na tabela das saídas (*Equal variances not assumed*).

O resultado foi  $t = 4786,448$  valor  $p \approx 0$  mostra que é possível rejeitar a hipótese  $H_0$ , ou seja, as médias diferem. Além disso, é fornecido que a diferença entre as médias desses dois países (Brasil e México) foi de 5,34 alunos, podendo variar com 95% de confiança entre 4,7 e 5,9 alunos. Ou seja, o México tem em torno de 5 alunos a mais por turma do que o Brasil.

O mesmo ocorre em relação à comparação por intervalo de confiança. O resultado fornecido não inclui o valor zero, possibilitando rejeitar  $H_0$  e concluir que as médias diferem ao nível de significância de 5%, já que o intervalo foi construído fixando o nível de significância de 5% (ou equivalentemente, um nível de confiança de 95%).

No terceiro exemplo, para testar a possibilidade de comparação de médias entre Espanha e Itália, as hipóteses testadas foram:  $H_0$ : A média de alunos por turma da Espanha *não difere* da média de alunos por turma da Itália; e  $H_1$ : A média de alunos por turma da Espanha *difere* da média de alunos por turma da Itália.

Tabela 17: DESCRIÇÃO DOS DADOS RELACIONADOS À COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE ESPANHA E ITÁLIA, PARA O TESTE PARAMÉTRICO.

| Group Statistics           |         |      |       |                |                 |
|----------------------------|---------|------|-------|----------------|-----------------|
|                            | países  | N    | Mean  | Std. Deviation | Std. Error Mean |
| Número de alunos por turma | Espanha | 3263 | 21,75 | 7,696          | ,135            |
|                            | Itália  | 4936 | 21,52 | 4,879          | ,069            |

Fonte: Base de dados da pesquisa TALIS, OCDE 2009.

A Tabela 17 apresenta as medidas descritivas dos dados, identificando uma amostra para a Espanha de 3.363 professores, com média de 22 alunos por turma e

desvio padrão de 8 alunos por turma, o erro padrão amostral é de 0,14. A Itália possui uma amostra de 4.936 professores, com uma média de 22 alunos por turma, e desvio padrão de 5 alunos por turma, com um erro padrão amostral de 0,7. Pode-se notar que, para ambos os países o desvio padrão da amostra é menor, ou seja, a dispersão entre as escolas pesquisadas, nos dois países, está mais próxima da média amostral.

Tabela 18: ANÁLISE DE COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE ESPANHA E ITÁLIA, PELO TESTE PARAMÉTRICO.

| Independent Samples Test   |                             |   |      |                              |          |                 |                 |                       |   |       |
|----------------------------|-----------------------------|---|------|------------------------------|----------|-----------------|-----------------|-----------------------|---|-------|
|                            |                             | Levene's Test for Equality of Variances |      | t-test for Equality of Means |          |                 |                 |                       |   |       |
|                            |                             | F                                       | Sig. | t                            | df       | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference |       |
|                            |                             |   |      |                              |          |                 |                 |                       | Lower                                     | Upper |
| Número de alunos por turma | Equal variances assumed     | 810,797                                 | ,000 | 1,662                        | 8197     | ,097            | ,231            | ,139                  | -,041                                     | ,503  |
|                            | Equal variances not assumed |   |      | 1,523                        | 4992,597 | ,128            | ,231            | ,152                  | -,066                                     | ,528  |

Fonte: Base de dados da pesquisa TALIS, OCDE 2009.

A Tabela 18 diz respeito à comparação de médias calculadas pelo teste paramétrico t de Student e a comparação de médias por intervalos. O teste de Levene para comparar se as variâncias foi de  $F = 810.787$  com valor  $p \approx 0$ , o que significa que não é possível concluir que as variâncias sejam homogêneas, ou seja, rejeita-se a hipótese das variâncias serem iguais. Sendo assim, o resultado do teste t que será utilizado é o da 2ª linha na tabela das saídas (*Equal variances not assumed*).

O resultado do teste foi de  $t = 1,523$  valor  $p = 0,128$ , mostrando que não é possível rejeitar a hipótese  $H_0$ , ou seja, mesmo com um tamanho de amostra diferente, pode-se concluir que as médias de alunos por turma na Espanha e na Itália não diferem estatisticamente. Além disso, analisando o intervalo de confiança, pode-se observar que este contém o valor 0, mostrando que não é possível afirmar que existe diferença estatística entre as médias desses dois países.

A seguir serão apresentados exemplos de comparação com base em testes não paramétricos.



### 3.3.2. Comparação com base em testes não paramétricos entre duas amostras independentes

A diferença entre os testes paramétricos e não paramétricos, é que o primeiro fornece as diferenças entre as médias e os intervalos, enquanto que no segundo, é baseado em classificação por postos.

Para demonstração do teste de Mann-Whitney ou Teste U (não paramétrico), foram utilizadas duas amostras independentes extraídas do banco de dados da pesquisa TALIS, nos mesmos exemplos utilizados para os testes paramétricos: comparação entre Brasil e Espanha; Brasil e México; e Espanha e Itália. Este teste também é feito por meio do SPSS 13.0.

Para comparação de médias entre Brasil e Espanha, também foram testadas as mesmas hipóteses:  $H_0$ : A média de alunos por turma do Brasil *não difere* da média de alunos por turma da Espanha;  $H_1$ : A média de alunos por turma do Brasil *difere* da média de alunos por turma da Espanha.

Tabela 19: DESCRIÇÃO DOS DADOS RELACIONADOS À COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE BRASIL E ESPANHA PARA O TESTE NÃO PARAMÉTRICO.

| Ranks                      |         |      |           |              |
|----------------------------|---------|------|-----------|--------------|
|                            | País    | N    | Mean Rank | Sum of Ranks |
| Número de alunos por turma | Brasil  | 5491 | 5527,13   | 30349449,50  |
|                            | Espanha | 3263 | 2442,90   | 7971185,50   |
|                            | Total   | 8754 |           |              |

Fonte: Base de dados da Pesquisa TALIS, OCDE 2009.

A Tabela 19 mostra a descrição dos dados para comparação de médias utilizando o teste não paramétrico de Mann-whitney, com a amostra de 5.491 professores brasileiros e a amostra de 3.263 professores espanhóis. Esse teste cria um *ranking* (ordenação) de todos os casos e depois compara estes *rankings* com os *rankings* de cada país. Assim, a referida tabela apresenta uma coluna com a classificação média (Mean Rank) para o Brasil de 5527,13 e de 2442,90 para a Espanha; e uma coluna com a soma dos postos (Sum of Ranks) de cada país, que no Brasil é de 30349449,50 e na Espanha é de 7971185,50. A classificação média é

obtida mediante a divisão da soma dos postos pelo número da amostra de cada país.

Aplicando o teste U por meio do SPSS 13.0, é possível testar a validade das hipóteses  $H_0$  e  $H_1$ . Assim, são colocados os valores observados na tabela 19 no programa e através do teste é apresentado o valor p, como mostra a (Tabela 20).

Tabela 20: ANÁLISE DA COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE BRASIL E ESPANHA PELO TESTE NÃO PARAMÉTRICO.

| Test Statistics <sup>b</sup>   | Número de alunos por turma |
|--|----------------------------|
| Mann-Whitney U   | 2645969,5                  |
| <i>Asymptotic</i> Significance <sup>28</sup> . (2-tailed <sup>29</sup> ) | 0,000                      |

b: Grouping variable: País

Fonte: Base de dados da Pesquisa TALIS, OCDE 2009.

A Tabela 20 mostra a análise da comparação das médias entre os países, cujo resultado do teste foi  $U = 2645969,5$  e valor  $p \approx 0$ , que mostra que é possível rejeitar a hipótese  $H_0$ . Assim, pode-se concluir que as médias de alunos por turma no Brasil e na Espanha são diferentes, porque o valor p é menor do que 0,05.

No segundo exemplo, as hipóteses testadas para comparação de médias entre Brasil e México foram:  $H_0$ : A média de alunos por turma do Brasil *não difere* da média de alunos por turma do México; e  $H_1$ : A média de alunos por turma do Brasil *difere* da média de alunos por turma do México.

Tabela 21: DESCRIÇÃO DOS DADOS RELACIONADOS À COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE BRASIL E MÉXICO PARA O TESTE NÃO PARAMÉTRICO.

| Ranks                      |        |      |           |              |
|----------------------------|--------|------|-----------|--------------|
|                            | País   | N    | Mean Rank | Sum of Ranks |
| Número de alunos por turma | Brasil | 5491 | 3947,07   | 21673336,00  |
|                            | México | 3101 | 4915,25   | 15242192,00  |
|                            | Total  | 8592 |           |              |

Fonte: Base de dados da Pesquisa TALIS, OCDE 2009.

<sup>28</sup> *Asymptotic* Significance: conhecido como valor p.

<sup>29</sup> 2-tailed: teste bilateral.

A Tabela 21 mostra a descrição dos dados para comparação de médias utilizando o teste não paramétrico de Mann-whitney, com amostra de 5.491 professores brasileiros, e 3.101 professores mexicanos, mostrando os rankings.

Tabela 22: ANÁLISE DA COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE BRASIL E MÉXICO PELO TESTE NÃO PARAMÉTRICO.

| Test Statistics <sup>b</sup> | Número de alunos por turma |
|------------------------------|----------------------------|
| Mann-Whitney U               | 6595050,0                  |
| Asymp. Sig. (2-tailed)       | 0,000                      |

b: Groupingvariable: País

Fonte: Base de dados da Pesquisa TALIS, OCDE 2009.

A Tabela 22 mostra a análise da comparação das médias entre os países, e o resultado do teste foi  $U = 6595050,0$  valor  $p \approx 0$  e mostra que é possível rejeitar a hipótese  $H_0$ . Assim, pode-se concluir que as médias do número de alunos por turma entre Brasil e México são diferentes com valor  $p > 0,05$ .

Esses testes mostram a comparação entre duas amostras independentes. Porém, também é possível analisar a diferença de médias entre vários grupos ao mesmo tempo. Como a pesquisa TALIS inclui vários países com diferentes tipos de amostras, também pode ser útil para testar os mesmos países selecionados de uma única vez.

### 3.3.3. Comparação entre três ou mais amostras independentes

Para exemplificar o funcionamento dos testes de múltiplas amostras ou *k amostras independentes* foi acrescentado mais um país para analisar a diferença entre quatro países, simultaneamente. Cabe ressaltar que poder-se-ia fazer esta análise com todos os países, ou seja, os testes de múltiplas amostras ou *k amostras independentes* não apresentam limites de amostras.

Assim, foram comparadas as amostras de Brasil, Espanha, México e Itália, com base em um teste paramétrico, denominado Teste Tukey; e o teste não paramétrico de Kruskal-Wallis. Foram testadas as mesmas hipóteses:  $H_0$ : A média de alunos por turma entre os países *não difere*; e  $H_1$ : pelo menos uma das médias de alunos por turma entre os países *difere*.

A Tabela 24 mostra a descrição dos dados referente aos países como já explicado nas análises anteriores, com um acréscimo dos limites mínimo e máximo, onde no Brasil, na Espanha e na Itália, a pesquisa mostrou que existem escolas com no mínimo 1 aluno por turma, e no máximo 98, 95 e 96 alunos por turma, respectivamente, e no México, com no mínimo 4 alunos e no máximo 98 alunos por turma.

Tabela 23: DESCRIÇÃO DOS DADOS RELACIONADOS À COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE OS PAÍSES, PARA O TESTE PARAMÉTRICO.

|         | N    | Mean  | Std. Deviation | Std. Error | 95% Confidence Interval for Mean |             | Minimum | Maximum |
|---------|------|-------|----------------|------------|----------------------------------|-------------|---------|---------|
|         |      |       |                |            | Lower Bound                      | Upper Bound |         |         |
| Brasil  | 5491 | 33,82 | 10,712         | ,145       | 33,54                            | 34,10       | 1       | 98      |
| Espanha | 3263 | 21,75 | 7,696          | ,135       | 21,48                            | 22,01       | 1       | 95      |
| México  | 3101 | 39,17 | 15,567         | ,280       | 38,62                            | 39,71       | 4       | 98      |
| Itália  | 4936 | 21,52 | 4,879          | ,069       | 21,38                            | 21,65       | 1       | 96      |

Fonte: Base de dados da Pesquisa TALIS, OCDE 2009.

A Tabela 23 analisa se pelo menos uma das variâncias difere entre os quatro países. Para tanto, é necessário fazer o teste da ANOVA (análise de variâncias) para testar se uma ou mais médias diferem entre si.

Tabela 24: TESTE DA VARIÂNCIA ANOVA, ENTRE OS PAÍSES PARA O TESTE PARAMÉTRICO.

|                | Sum of Squares | df    | Mean Square | F        | Sig. |
|----------------|----------------|-------|-------------|----------|------|
| Between Groups | 895797,47      | 3     | 298599,158  | 2962,836 | ,000 |
| Within Groups  | 1691819,8      | 16787 | 100,782     |          |      |
| Total          | 2587617,2      | 16790 |             |          |      |

Fonte: Base de dados da Pesquisa TALIS, OCDE 2009.

A Tabela 24 da análise de variância apresenta a variabilidade entre os países (*between groups*) e o grupo residual (*within groups*). O teste  $F=2962.836$  e valor  $p \approx 0$  conclui que existem diferenças significativas entre a média de alunos por turma entre pelo menos dois dos países, com um nível de significância de 5%. Contudo, este resultado não permite concluir qual ou quais países diferem na média de alunos por turma. Então, faz-se necessária a comparação da média de alunos por turma através do teste de Tukey, para identificar se um país ou mais de um difere dos

demais. Este é o princípio de qualquer teste de comparações múltiplas, utilizando as hipóteses  $H_0$  e  $H_1$ , citadas anteriormente.

Tabela 25: COMPARAÇÃO MÚLTIPLA ENTRE AS MÉDIAS DE ALUNOS POR TURMA ENTRE BRASIL (1), ESPANHA (2), MÉXICO (3) E ITÁLIA (4), PELO TESTE DE TUKEY.

| Multiple Comparisons         |              |                       |            |      |                         |             |
|------------------------------|--------------|-----------------------|------------|------|-------------------------|-------------|
| Dependent Variable: VAR00002 |              |                       |            |      |                         |             |
| Tukey HSD                    |              |                       |            |      |                         |             |
| (I) VAR00001                 | (J) VAR00001 | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval |             |
|                              |              |                       |            |      | Lower Bound             | Upper Bound |
| 1                            | 1            |                       |            |      |                         |             |
|                              | 2            | 12,074*               | ,222       | ,000 | 11,50                   | 12,64       |
|                              | 3            | -5,345*               | ,226       | ,000 | -5,92                   | -4,77       |
|                              | 4            | 12,305*               | ,197       | ,000 | 11,80                   | 12,81       |
| 2                            | 1            | -12,074*              | ,222       | ,000 | -12,64                  | -11,50      |
|                              | 2            |                       |            |      |                         |             |
|                              | 3            | -17,419*              | ,252       | ,000 | -18,07                  | -16,77      |
|                              | 4            | ,231                  | ,227       | ,738 | -,35                    | ,81         |
| 3                            | 1            | 5,345*                | ,226       | ,000 | 4,77                    | 5,92        |
|                              | 2            | 17,419*               | ,252       | ,000 | 16,77                   | 18,07       |
|                              | 3            |                       |            |      |                         |             |
|                              | 4            | 17,650*               | ,230       | ,000 | 17,06                   | 18,24       |
| 4                            | 1            | -12,305*              | ,197       | ,000 | -12,81                  | -11,80      |
|                              | 2            | -,231                 | ,227       | ,738 | -,81                    | ,35         |
|                              | 3            | -17,650*              | ,230       | ,000 | -18,24                  | -17,06      |
|                              | 4            |                       |            |      |                         |             |

\*. The mean difference is significant at the .05 level.

Fonte: Base de dados da Pesquisa TALIS, OCDE 2009.

Na Tabela 25, os resultados dos testes de comparações múltiplas são identificados com asterisco(\*) para identificar os tratamentos cujas médias de alunos por turma diferem significativamente, no caso, estabelecendo confronto entre Brasil, Espanha e México ou Brasil, Itália e México. Onde não há asterisco, nos resultados, significa que as médias das variâncias de alunos por turma não diferem, no caso, entre Espanha e Itália, com valor  $p = 0,738$ , ou seja, superior ao nível de significância de 0,05%.

Para exemplificar a análise do teste não paramétrico Kruskal-Wallis foram utilizadas as mesmas hipóteses e países com as comparações de médias de alunos por turma, e os resultados são mostrados na Tabela 26.

Tabela 26: DESCRIÇÃO DOS RANKS RELACIONADOS À COMPARAÇÃO DE MÉDIAS DE ALUNOS POR TURMA ENTRE OS PAÍSES.

|                            | Ranks   |       |          |
|----------------------------|---------|-------|----------|
|                            | Países  | N     | MeanRank |
| Número de alunos por turma | Brasil  | 5491  | 11107,13 |
|                            | Espanha | 3263  | 5481,66  |
|                            | México  | 3101  | 12196,29 |
|                            | Itália  | 4936  | 4919,08  |
|                            | Total   | 16791 |          |

Fonte: Base de dados da Pesquisa TALIS, OCDE 2009.

A Tabela 26 mostra os levantamentos descritivos dos *ranks* médios das médias de alunos por turma. E a Tabela 28 mostra se existe diferença entre pelo menos uma das médias de alunos entre os países.

A Tabela 27 conclui, da mesma forma que realizados nos testes de Tukey no procedimento paramétrico, com um grau de significância de 0,05% e valor  $p \approx 0$  que pelo menos uma das variâncias difere<sup>30</sup>.

Tabela 27: ANÁLISE DE TESTE KRUSKAL-WALLIS, ANALISANDO SE PELO MENOS UMA DAS MÉDIAS DE ALUNOS POR TURMA DIFERE ENTRE OS PAÍSES.

| Test Statistics <sup>a,b</sup> |         |
|--------------------------------|---------|
| Número de alunos por turma     |         |
| Chi-Square                     | 7361,14 |
| df <sup>31</sup>               | 3       |
| Asymptotic Significance        | 0,000   |

a: Kruskal-Wallis test

b: Grouping Variable: Países

Fonte: Base de dados da Pesquisa TALIS, OCDE 2009.

<sup>30</sup> Para identificar qual país ou quais países diferem entre si pode ser utilizado o método de estatística experimental não paramétrica de Campos (1983). Contudo, esse procedimento é pouco conhecido e não será utilizado neste estudo. Ver: CAMPOS, H. Estatística Experimental Não-Paramétrica. São Paulo: ESALQ, 1983.

<sup>31</sup> df: número dos graus de liberdade.

O teste não paramétrico deve ser utilizado quando as amostras forem pequenas e a dispersão for grande, havendo heterogeneidade entre os dados. No exemplo utilizado, da pesquisa TALIS, basta a aplicação do teste paramétrico para testar a comparação, porque as amostras são suficientemente grandes e a mediana não difere significativamente da média.

Desse modo, os testes aplicados permitem afirmar que é possível estabelecer, estatisticamente, que se a pesquisa TALIS fosse aplicada, nas mesmas condições, para todos os professores do ensino secundário básico desses dois países apresentaria a mesma diferença encontrada na amostra – em torno de 5 alunos por turma. Da mesma forma, se coloca a mesma possibilidade para a comparação entre Brasil e Espanha – em torno de 12 alunos por turma – e entre Espanha e Itália – que apresentam médias semelhantes.

De tal modo, justifica-se que, para evitar erros na realização de comparações, é necessário compreender e aplicar, adequadamente, os testes inferenciais, aqui demonstrados, para ser possível uma conclusão adequada e ultrapassar a mera descrição.

## Considerações Finais

No que concerne ao emprego dos métodos quantitativos, GATTI (2004) adverte que é necessário considerar dois aspectos importantes: primeiro, em se tratando de dados descritivos (números, frequências, escalas e medidas), que se pode, por vezes, não alcançar o objetivo esperado, porque em muitos casos os questionários são construídos sem um propósito específico; segundo, que boas análises dependem de boas perguntas com propósitos para um bom levantamento de dados, com um referencial teórico de qualidade e da perspectiva epistêmica na abordagem do problema de pesquisa.

Sem considerar as condições adequadas, indicadas por Gatti, corre-se o risco de usar certos tratamentos estatísticos indevidamente e interpretações equivocadas. A estatística informa os números, porém a interpretação desses resultados é dada pelo pesquisador, em função de seu entendimento teórico. Por isso, é necessário o domínio do conhecimento estatístico para obter conclusões mais precisas sobre os acontecimentos investigados.

No atual mundo globalizado e informatizado, esta ferramenta cresceu e ampliou seu espaço de influência. Entretanto, o número de profissionais formados não acompanha o aumento da demanda. Atualmente, o Brasil conta com 38 cursos de graduação em Estatística reconhecidos pelo MEC (BRASIL, 2014a) e 17 cursos de pós-graduação *stricto sensu*, 10 de mestrado e 7 de doutorado, também reconhecidos (BRASIL, 2014b). Mas, a procura por esse tipo de curso no país ainda é muito baixa, apesar de sobrar vagas no mercado de trabalho e das excelentes perspectivas de carreira e salário.

Entretanto, não se deve esquecer que a estatística é apenas um instrumento a serviço do pesquisador e nem sempre capaz de fornecer resultados conclusivos, como é o caso das pesquisas realizadas por organizações internacionais e órgãos oficiais de governo. Mas, nem por isso deve-se rejeitá-la, tendo em vista que esta não cria informação que não existe, pois seu grande mérito é evidenciar fatos desconhecidos, assim como melhorar os mecanismos já aplicados, destacando o que realmente é essencial para facilitar as tomadas de decisão.

No âmbito das reformas de Estado dos anos 1990, a educação básica passou a ser vista como alvo principal para o desenvolvimento social e econômico e os sistemas de avaliação foi criado para avaliar a adequação ao modelo estabelecido e,



deste modo, a generalização global desse modelo reduz a importância da avaliação que é feita pelo professor na sala de aula e fortalece o monitoramento externo e centralizado. O que pode significar uma inversão arriscada, onde a avaliação passa a determinar o processo de execução das políticas.

Para que a avaliação possa subsidiar tomada de decisões políticas, deve ser tomada com um processo de análise e acompanhamento da implementação de políticas, que inclua o diagnóstico, o planejamento, e a execução e seus resultados. Assim, para que seja efetiva, a avaliação não pode ser realizada de qualquer maneira. É necessário avaliar um determinado momento – avaliação estática, que corresponde ao final de um processo – mas, também o processo e o movimento histórico de desenvolvimento dos sistemas – avaliação dinâmica.

No caso da pesquisa TALIS, pode-se afirmar que a análise realizada pelo relatório (OCDE, 2009) apresenta uma avaliação estática sobre a formação e as condições de trabalho dos professores investigados, pois se limita à descrição dos resultados obtidos por apenas um instrumento, sem considerar o processo e o contexto ao qual se refere. Neste sentido, pode-se afirmar que nas pesquisas de larga escala, realizadas em âmbito nacional e internacional, não cabe uma avaliação dinâmica, tendo em vista que estas não incluem estudo sobre os processos e os contextos relacionados com o objeto da avaliação.

O processo de elaboração de questões a partir de indicadores é sempre complexo e exige um conhecimento especializado da área da Estatística, que a maioria dos profissionais de outras áreas, como também os da educação, desconhece. Tal desconhecimento pode levar a uma leitura superficial e, por vezes, equivocada dos resultados apresentados em relatórios que, ao final, vão orientar a avaliação e a proposição de políticas.

A definição de indicadores expressarem, claramente, o que se pretende, ou neste caso, as finalidades da educação. Assim, para criar um indicador é necessário estabelecer um conjunto de critérios de avaliação que se relacionem, objetivamente, com a concepção de qualidade de um determinado objeto.

As pesquisas de grande porte são promovidas e financiadas por agências ou órgãos governamentais que contratam profissionais empresas privadas para desenvolver todos os procedimentos técnicos necessários de acordo com os seus interesses. Nesse contexto, a educação passou a ser *presa fácil* de estudos

internacionais, que estabelecem seus padrões de qualidade, sem resistência dos profissionais da educação.

Como nas pesquisas de larga escala é muito difícil obter dados de toda população, a estratégia é selecionar uma amostra representativa para, a partir desta, estimar como é a população. Contudo, a definição da amostra deve ser rigorosa para que os resultados se aproximem ao máximo da realidade a ser observada, assegurando a possibilidade de calcular uma margem segura de confiabilidade.

Para obter essa margem de segurança a Estatística oferece métodos e técnicas que possibilitam provar ou refutar hipóteses em um dado estudo. Nesse contexto, torna-se imperativo o auxílio de pessoas habilitadas em conhecimento estatístico para planejar, coletar, organizar e interpretar corretamente os dados necessários para realização de uma pesquisa.

A estatística descritiva compreende conceitos básicos para determinar a forma que o estudo deve ser seguido por meio de levantamentos de dados de uma análise numérica. É através desse estudo que se observam possíveis tendências, erros de digitação, desenvolvem-se tabelas e gráficos e identifica-se o tipo da variável (numérica, ordinal, intervalar ou de razão) com a qual se irá trabalhar. Mas, este ramo da estatística não informa o resultado, apenas os descreve.

Nesse tipo de estatística, o mais importante é a *distribuição de frequência*. Ou seja, os dados são arrumados em números distribuídos em tabelas e, através destas, podem ser criados gráficos, identificadas como *medidas de tendência central* e de *dispersão*.

As medidas de tendência central são indicadores que permitem que se tenha uma ideia ou um resumo, de um determinado conjunto de dados, mostrando apenas um valor específico. E a dispersão são valores que estão em torno da medida de tendência central, identificando a distância de cada valor.

O uso apropriado dessas medidas pode evidenciar tendências mais próximas à realidade. De modo geral, as pesquisas educacionais divulgam apenas as médias que, nem sempre são as medidas mais adequadas para compreender determinado problema, e o desvio padrão, que mostra o grau de homogeneidade ou heterogeneidade da referida média, é ignorado.

Em estudos com coleta de dados amostrais, que representam uma população, as amostras estão associadas a uma margem de erro, sendo, portanto,

de fundamental importância calcular suas estimativas. Para isso, faz-se necessária a estatística inferencial, que trabalha com as análises dos dados preparados pela estatística descritiva, inferindo se a hipótese testada deve ser rejeitada ou não. Para tanto, a estatística permite extrair dos dados as informações necessárias com base em intervalos de confiança e redução da margem de erro.

Em todos os testes inferenciais, as decisões acerca das características da população baseiam-se em hipóteses que, de um modo geral, são afirmações suportadas pelas distribuições dos dados relativos às características da amostra populacional que se pretendem testar. Para tanto é necessário considerar a *hipótese nula* ( $H_0$ ) e a *hipótese alternativa* ( $H_1$ ), tanto nos testes paramétricos quanto nos não paramétricos, através dos softwares estatísticos – como o SPSS, por exemplo – que expressam a significância estatística através do valor  $p$ , que permite decidir sobre a validade ou não da hipótese nula.

Teoricamente, os testes paramétricos são os melhores e mais poderosos para avaliar comparações. Porém, na prática, nem sempre acontece dos dados estarem dentro dos pré-requisitos exigidos por esses testes. É muito comum encontrar questionários respondidos erradamente, faltando dados, amostras muito pequenas, respostas com valores absurdos (muito conhecidos na estatística como *outliers*) e variáveis categoricamente mensuradas como nominais e ordinais. Para tanto, a estatística fornece os testes não paramétricos para trabalhar com esses dados, a fim de chegar a resultados mais precisos.

Deste modo, a estatística inferencial, considerando o erro de estimação, permite dimensionar a confiança que se pode ter no resultado de uma pesquisa ou então subsidiar a tomada de decisão, para evitar a aceitação ou rejeição de uma hipótese estatística.

Mas, antes de tudo, deve-se considerar a importância da elaboração dos questionários para a pesquisa, assim como das condições oferecidas aos respondentes. Os questionários da pesquisa TALIS foram enviados aos professores e diretores, segundo a OCDE (2009), de forma impressa, ou *online*, o qual é denominado instrumento autoaplicável. Nos dois casos, deve oferecer toda a informação necessária para obter as respostas da maneira esperada pelo pesquisador. Mas, pelas informações obtidas por respondentes mexicanos, estas não foram suficientes.

Além disso, um questionário mal elaborado impede a aplicação adequada da estatística e pode gerar resultados equivocados. A determinação da escala inadequada em um questionário pode levar a resultados equivocados, conforme se observou no uso da Escala Likert de quatro pontos para medir as crenças, atitudes e práticas dos professores. Como a referida escala não inclui uma categoria neutra, pode-se deduzir que há possibilidade de haver indução de determinada tendência naquelas questões.

Outros problemas que precisam ser destacados, analisando as tabelas descritivas apresentadas no relatório da OCDE 2009, se referem à indicação do erro padrão, em relação ao porte das escolas que fizeram parte da pesquisa e não do desvio padrão, que seria mais relevante para mostrar a dispersão em relação à média. Se for analisada somente a média – como é de costume na análise das pesquisas educacionais –, nota-se que o Brasil possui um número médio de 880 alunos matriculados por escola, assim como a Espanha possui 668 e o México 554. Nesta perspectiva, pode-se levantar a hipótese de que as escolas que participaram da pesquisa possuem um porte relativamente alto. Mas, quando confrontado com o desvio padrão, observa-se que existe uma dispersão extremamente alta em torno da média, isto é, existe uma grande heterogeneidade entre os tamanhos das escolas na pesquisa.

O coeficiente de variação, que indica o percentual de dispersão em torno da média de 72% para o Brasil, 59% para a Espanha e 76% para o México, evidencia o equívoco na utilização da média. Neste caso, deveria ter sido utilizada a mediana (ou outros percentis<sup>32</sup>) que apresenta uma diferença de quase 100 alunos em relação à média, nos casos do Brasil e do México e 68 alunos, no caso da Espanha. Este fato indica também a importância de se considerar o coeficiente de variação para interpretação dos dados, no caso de pesquisas que trabalham com diversas amostras, afim de identificar com mais precisão a dispersão em torno da média de cada país.

Em muitas pesquisas é comum que se tenha interesse na comparação de medidas de tendência central. Na pesquisa TALIS, por exemplo, pode-se desejar

---

<sup>32</sup> Percentil - Em estatística descritiva, os percentis são medidas que dividem a amostra ordenada (por ordem crescente) em 100 partes, cada uma com uma porcentagem de dados aproximadamente igual. O 1º percentil determina 1% menor dos dados e o 98º percentil, determina o 98% menor dos dados. O 25º percentil é o primeiro quartil; o 50º percentil é a mediana. Disponível: <http://wikiciencias.casadasciencias.org/wiki/index.php/Percentis>.

comparar se a média de alunos por turma obtida por dois países (Brasil e Espanha, ou Brasil e México, por exemplo) diferem estatisticamente entre si.

Como os valores obtidos para as médias são baseadas em amostras extraídas das populações de interesse, se faz necessário o uso de testes de comparação ou obtenção de intervalos de confiança a fim de, com base nas amostras observadas, permitir a realização de inferências a respeito de possíveis diferenças estatísticas entre essas medidas, para duas ou mais populações comparadas, por meio do programa SPSS 13.0, que faz uma análise das variâncias amostrais (dispersão), apresentando as saídas possíveis para identificar o tipo de teste que deve ser utilizado.

O teste não paramétrico deve ser utilizado quando as amostras forem pequenas e a dispersão for grande, havendo heterogeneidade entre os dados. No exemplo utilizado, da pesquisa TALIS, bastava a aplicação do teste paramétrico para testar a comparação, porque as amostras são suficientemente grandes e a mediana não difere significativamente da média.

Desse modo, os testes aplicados permitem afirmar que é possível estabelecer, estatisticamente, que se a pesquisa TALIS fosse aplicada, nas mesmas condições, para todos os professores do ensino secundário básico apresentaria a mesma diferença encontrada na amostra – em torno de 5 alunos para o Brasil e o México. Da mesma forma, se coloca a mesma possibilidade para a comparação entre Brasil e Espanha – em torno de 12 alunos por turma – e entre Espanha e Itália – que apresentam médias semelhantes.

De tal modo, justifica-se que, para evitar erros na realização de comparações, é necessário compreender e aplicar, adequadamente, os testes inferenciais, aqui demonstrados, para ser possível uma conclusão adequada e ultrapassar a mera descrição.

A partir das considerações descritas, pode-se avaliar a importância do uso da Estatística para subsidiar a avaliação e implementação de políticas educacionais. A pesquisa TALIS, tomada como referência, permitiu uma análise exemplar de problemas que podem ser causados nas conclusões tiradas a partir de questionários mal elaborados – como na questão sobre o barulho em sala de aula –, no número elevado de questões e itens – que pode ocasionar cansaço e levar a respostas irrefletidas –, na escolha inadequada de escalas – como no caso da Escala Likert de

quatro pontos –, no equívoco da identificação do desvio padrão nas tabelas – como a indicação de desvio padrão para o conteúdo referente a erro padrão –, entre outros fatores.

Porém, estas análises não permitem, nos limites de tempo disponibilizados para esta pesquisa, fazer uma análise conclusiva do tema proposto. Para aperfeiçoar, desenvolver e aprofundar as observações encontradas no presente estudo para avaliar o uso da Estatística, como meio de subsidiar a avaliação e implementação de políticas educacionais, ainda há muito a ser feito. No que se refere ao exemplo utilizado, qual seja o programa de pesquisa TALIS, sugere-se um estudo de caso, para replicar os questionários para professores e diretores, a fim de verificar como esses respondentes interpretam as questões formuladas, o grau de cansaço e estresse causado pela extensão de questões e itens e o entendimento das opções de resposta da Escala Likert aplicada, considerando o que foi aqui analisado.

## REFERÊNCIAS

Ação Educativa, Unicef, PNUD, Inep-MEC (coord.). Indicadores da qualidade na educação. São Paulo: **Ação Educativa**, 2004. Disponível em: [http://portal.mec.gov.br/seb/arquivos/pdf/Consescol/ce\\_indqua.pdf](http://portal.mec.gov.br/seb/arquivos/pdf/Consescol/ce_indqua.pdf). Acesso: 15/04/2014.

AFONSO, Almerindo Janela. Reforma do Estado e Políticas Educacionais: entre a crise do Estado-nação e a emergência da regulação supranacional. **Educação & Sociedade**. São Paulo, ano XXII, n. 75, agosto/2001, p. 15-32. Disponível em: <http://www.scielo.br/pdf/es/v22n75/22n75a03>. Acesso dia 20/08/2014.

ALVES, Thiago. Avaliação na Administração Pública. In: ALVES, Thiago; PASSADOR, Cláudia Souza. **Educação Pública no Brasil**: condições de oferta, nível socioeconômico dos alunos e avaliação. São Paulo: Annablum; Brasília: CAPES, INEP, 2011, p. 33-49.

AMARO, Ana; PÓVOA, Andreia; MACEDO, Lúcia. A arte de fazer questionário. **Metodologia de Investigação em Educação**, Faculdade de Ciências da universidade do Porto, Portugal. 2005. Disponível em: [http://www.unisc.br/portal/upload/com\\_arquivo/a\\_arte\\_de\\_fazer\\_questionario.pdf](http://www.unisc.br/portal/upload/com_arquivo/a_arte_de_fazer_questionario.pdf). Acesso: 23/01/2015.

ANDRADE, Ana Cristine Teixeira. **A liderança na escola e os professores** - Instituto Superior de Educação e Trabalho (ISET). Porto, julho de 2010. Disponível em: [http://www.iset.pt/iset/DissertacoesPDF/6\\_ana\\_cristina\\_andrade\\_web.pdf](http://www.iset.pt/iset/DissertacoesPDF/6_ana_cristina_andrade_web.pdf). Acesso: 03/04/2014.

AZEVEDO, Amílcar Gomes de; CAMPOS, Paulo Henrique Borges de . – Estatística Básica Cursos de Ciências Humanas e de Educação. 4 ed. Rio de Janeiro: Livros Técnicas e Científicos, 1981.

BABBIE, Earl. **Métodos de pesquisa de survey**. Trad. Guilherme Cezarino. Belo Horizonte: Ed. UFMG, 1999.

BANDEIRA, Marina. Texto11. Como elaborar um questionário. Laboratório de Pesquisa Ambiental. Série – Planejamento de Pesquisas nas Ciências Sociais. Universidade de Brasília. N.01. 2003. Disponível em: [http://www.ufsj.edu.br/portal2-repositorio/File/lapsam/Metodo%20de%20pesquisa/Metodos%20de%20pesquisa%202013/Texto\\_11-\\_Como\\_elaborar\\_um\\_quesitonario.pdf](http://www.ufsj.edu.br/portal2-repositorio/File/lapsam/Metodo%20de%20pesquisa/Metodos%20de%20pesquisa%202013/Texto_11-_Como_elaborar_um_quesitonario.pdf) .Acesso: 10/02/2015.

BANDURA, Albert. **Self-Efficacy**. In: Ramachaudran, V. S. (ed.), Encyclopedia of human behavior (Vol. 4, pp. 71-81). New York: AcademicPress, 15/02/1997. Disponível em: <http://www.westga.edu/~vickir/Healthcare/HC14%20FacititatingResiliency/Link%2010%20--%20Self-Efficacy.pdf> Acesso: 07/04/2014.

BAYER, Arno; BITTENCOURT, Hélio ; ROCHA, Josy ; ECHEVESTE, Simone. **Estatística e sua História**. 2004. Disponível em:

[http://www.exatas.net/ssbec\\_estatistica\\_e\\_sua\\_historia.pdf](http://www.exatas.net/ssbec_estatistica_e_sua_historia.pdf). Acesso em: 6/01/2013.

BELLO, José Luiz de Paiva, Educação no Brasil: a História das rupturas. **Pedagogia em Foco**, Rio de Janeiro, 2001. Disponível em: <http://www.pedagogiaemfoco.pro.br/heb14.htm>. Acesso em: 10/10/2012.

BÍBLIA. São Paulo: Ed. Ave Maria, 1997.

BOLFARINE, Helleno; BUSSAB Wilton O. **Elementos de Amostragem**: Ed. Edgard Blucher. 1 edição. 2005.

BRASIL a. Ministério da Educação. **Instituições de Educação Superior e Cursos Cadastrados**. Brasília, 2014. Disponível em: <https://emec.mec.gov.br/>. Acesso em: 29/08/2014.

BRASIL b. Coordenação e aperfeiçoamento de pessoal de nível superior. **Cursos recomendados e reconhecidos**. Brasília, 2014. Disponível em: <http://www.capes.gov.br/avaliacao/cursos-recomendados-e-reconhecidos>. Acesso: 29/08/2014.

BRASIL c, Ministério da Educação. Secretaria do Ensino Fundamental. **Parâmetros curriculares nacionais: Matemática**: MEC/SEF, Brasília, 1998. Disponível em: <http://portal.mec.gov.br/seb/arquivos/pdf/matematica.pdf> . Acesso 20/01/2015. Acesso: 02/02/2015.

BRITO, Marica Regina F. de: **O SINAES e o ENADE: Da Concepção à Implementação** , Avaliação, Campinas; Sorocaba, SP, v.13, n.3, p. 841-850, nov. 2008. Disponível em: <http://www.scielo.br/pdf/aval/v13n3/14.pdf>. Acesso: 06/02/2015

CHAGAS, Anivaldo Tadeu Roston: O questionário da pesquisa – Administração Online Prática – Pesquisa – Ensino. V.1. n. 1. Jan/fev/mar 2000. Disponível em: [http://www.fecap.br/adm\\_online/art11/anival.htm](http://www.fecap.br/adm_online/art11/anival.htm). Acesso: 07/02/2015.

CALDAS, Andreia do Rocio. **Desistência e resistência no Trabalho Docente**: um estudo das professoras e professores do Ensino Fundamental da Rede Municipal de Educação de Curitiba (2007). Tese de Doutorado. (Educação). Curitiba: UFPR.

CARISSIMI, Aline ChalusVernick; TROJAN, Rose Meri. A valorização do professor no Brasil no contexto das tendências globais. In: **Jornal de Políticas Educacionais/Núcleo de Políticas, Gestão e Financiamento da Educação da Universidade Federal do Paraná –NUPE/UFPR** – v. 1, n.1 (1.º semestre de 2007) – Curitiba: NuPE/UFPR, volume 5, número 10 – Agosto/Dezembro de 2011.

CASASSUS, Juan. A reforma da América Latina no contexto da globalização. In: Cadernos de pesquisa.N.114, p.7-28, Nov. 2001. Disponível em: <http://www.scielo.br/pdf/cp/n114/a01n114.pdf>. Acesso em: 16/02/2014.

CASTENHEIRA, Nelson Pereira – Estatística Aplicada a todos os níveis – Ed. IBPEX Diológica, 5 ed. Curitiba 2010.



CRESPO, Antonio Arnot. **Estatística Fácil**: São Paulo, Ed. Saraiva - 18ª edição – 2002.

CORRADINI, Suely Nercessian. **Indicadores de Qualidade na Educação: um estudo a partir do PISA e da TALIS**. Teses (Doutorado). Biblioteca Comunitária. UFscar. São Paulo. 2012. Disponível em: [http://www.bdt.d.ufscar.br/htdocs/tedeSimplificado//tde\\_busca/arquivo.php?codArquivo=5562](http://www.bdt.d.ufscar.br/htdocs/tedeSimplificado//tde_busca/arquivo.php?codArquivo=5562). Acesso dia: 10/04/2014.

CORRALES, Javier. **Aspectos Políticos na implementação das reformas educacionais**. Traduzido por: Paulo Martins Garchet, PREAL, P. 01-62, Massachusetts, EUA, abril 2000. Disponível em: [http://www.oei.es/reformaseducativas/aspectos\\_politicos\\_implementacion\\_reformas\\_educativas\\_corrales\\_portugues.pdf](http://www.oei.es/reformaseducativas/aspectos_politicos_implementacion_reformas_educativas_corrales_portugues.pdf). Acesso: 16/05/2014.

DALE, Roger. Globalización: ¿un nuevo mundo para la educación comparada? In: SCHRIEWER, Jürgen (org.) **Formación del discurso en la educación comparada**. Barcelona: Pomares, 2002, p.69-90.

\_\_\_\_\_. **Globalização e Educação**: demonstrando a existência de uma “cultura educacional mundial comum” ou localizando uma “agenda globalmente estruturada para a educação”? – Educ. Soc. Campinas, vol. 25, n. 87, p. 423-460, maio/ago. 2004. Disponível em: <<http://www.cedes.unicamp.br>> Acesso: 28/08/2014.

D'HAINAUT, Louis: Conceitos e Métodos da Estatística – **Uma Variável a uma Dimensão**. Vol 1. 1990. Lisboa: Fundação Calouste Gulbenkian

DIAS SOBRINHO. José. **Avaliação e Transformações da Educação Superior Brasileira (1995-2009): Do PROVÃO ao SINAES**, Avaliação, Campinas; Sorocaba, SP, v. 15, n. 1, p. 195-224, mar. 2010. Disponível em: <http://www.scielo.br/pdf/aval/v15n1/v15n1a11.pdf>. Acesso: 05/02/2015.

DIÁRIO DE CUIABÁ. **Editorial**. Plano Nacional da Educação. Disponível em: <http://www.diariodecuiaba.com.br/detalhe.php?cod=419570>

DIEGUEZ, José David Alarid. El magisterio mexicano y el SNTE ante las transformaciones del trabajo docente en México. In: **Jornal de Políticas Educacionais/Núcleo de Políticas, Gestão e financiamento da Educação da Universidade Federal do Paraná –NUPE/UFPR** – v. 1, n.1 (1.º semestre de 2007) – Curitiba: NuPE/UFPR, volume 5, número 10 – Agosto/Dezembro de 2011.

FARIAS, Ana Maria Lima de. **Inferência Estatística**. Rio de Janeiro, Universidade Federal Fluminense, 2008. (Tese de doutorado). Disponível em: <http://www.professores.uff.br/anafarias/inferencia.pdf>. Acesso: 17/10/ 2014.

FERREIRA, Eliza Bartolozzi; OLIVEIRA, Dalila Andrade.(Orgs). **Crise da Escola e Políticas Educativas** . Belo Horizonte: Autêntica, 2009

FREITAS, Luiz Carlos de. **Eliminação adiada**: o ocaso das classes populares no interior da escola e a ocultação da (má) qualidade no ensino (2007). Educação &

Sociedade, Volume 28, Edição100, Páginas965-987. Disponível em:  
<http://www.redalyc.org/pdf/873/87313704016.pdf>. Acesso: 25/08/2014

GATTI, Bernadete: Estudos Quantitativos em Educação – In. **Revista Educação e Pesquisa**, São Paulo. V.30. n.1, p11-30, jan/abr. 2004. Disponível em:  
<http://www.scielo.br/pdf/ep/v30n1/a02v30n1.pdf>. Acesso: 05/02/1015.

GIL, Antônio Carlos. **Como elaborar projetos de Pesquisa**. 5ª. ed. São Paulo: Atlas, 2010. 184 p.

GOODE, Wilson; HATT, Paul K. **Métodos em pesquisa social**. 4ª ed. São Paulo: Nacional, 1973. xi. 488 p.

INEE. Instituto Nacional para la Evaluación de la Educación. **Ley del Instituto Nacional para la Evaluación de la Educación**. México: 2013. Disponível em:  
[http://www.inee.edu.mx/images/stories/2013/nueva\\_version/LeyINEEweb\\_ok.pdf](http://www.inee.edu.mx/images/stories/2013/nueva_version/LeyINEEweb_ok.pdf)  
 Acesso: 20/05/2014.

INEP. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira **Pesquisa TALIS**. Brasília: 2011. Disponível em: <http://portal.inep.gov.br/web/talis>  
 Acesso: 25/10/2012.

\_\_\_\_INEP. Instituto Nacional de Pesquisas Internacionais Anísio Teixeira. **Pisa**. Disponível em: <http://portal.inep.gov.br/pisa-programa-internacional-de-avaliacao-de-alunos> Acesso: 15/04/2014.

INE. Instituto Nacional de Estadística (INE) - **Algunos Antecedentes sobre la Organización para la cooperación y desarrollo económico** –Subdirección Técnica: Departamento de Metodología Estadística, Proyecto: Marco Metodológico de la Producción Estadística de Calidad. Enero 2006. Disponível em:  
[http://www.bcn.cl/carpeta\\_temas\\_profundidad/copy3\\_of\\_temas\\_profundidad.2007-05-02.5434448168/documentos\\_pdf.2007-06-28.4716180007/archivos\\_pdf.2007-06-28.5450664703/carpeta\\_temas\\_profundidad/temas\\_profundidad.2007-07-25.4772415999/documentos\\_pdf.2007-06-28.4716180007/archivos\\_pdf.2007-06-28.5450664703/archivo1](http://www.bcn.cl/carpeta_temas_profundidad/copy3_of_temas_profundidad.2007-05-02.5434448168/documentos_pdf.2007-06-28.4716180007/archivos_pdf.2007-06-28.5450664703/carpeta_temas_profundidad/temas_profundidad.2007-07-25.4772415999/documentos_pdf.2007-06-28.4716180007/archivos_pdf.2007-06-28.5450664703/archivo1). Acesso 27/08/2014

IRIS. Improvement Through Research in the Inclusive School.**Clima da sala de aula em ambientes inclusivos**. Education and Culture Lifelong Learning Programme Comenius, 2006. Disponível em:  
[http://www.irisproject.eu/teachersweb/PT/docs/TT\\_Clima\\_Sala\\_de\\_Aula\\_WD\\_PT.pdf](http://www.irisproject.eu/teachersweb/PT/docs/TT_Clima_Sala_de_Aula_WD_PT.pdf)  
 Acesso: 30/04/2014.

IRIS. Improvement Through Research in the Inclusive School.**Clima da sala de aula em ambientes inclusivos**. Education and Culture Lifelong Learning Programme Comenius, 2006. Disponível em:  
[http://www.irisproject.eu/teachersweb/PT/docs/TT\\_Clima\\_Sala\\_de\\_Aula\\_WD\\_PT.pdf](http://www.irisproject.eu/teachersweb/PT/docs/TT_Clima_Sala_de_Aula_WD_PT.pdf)  
 Acesso: 30/04/2014.

KLEIN, Ruben; FONTANIVE, NilmaSantos. **Avaliação em Larga Escala: uma proposta inovadora**. Em Aberto, Brasília, ano 15, n.66, abr/jun. 1995. Disponível

em: [rbep.inep.gov.br/index.php/emaberto/article/viewFile/995/899](http://rbep.inep.gov.br/index.php/emaberto/article/viewFile/995/899). Acesso 06/02/2015.

KRAWCZYK, Nora Rut; VIEIRA, Vera Lucia. **A reforma educacional no México e no Chile: apontamentos sobre as rupturas e continuidades**. *Educ. Rev.* [online] 2003, n.22. Disponível em: <http://www.redalyc.org/pdf/1550/155017964005.pdf>. Acesso: 15/02/2014.

LAFLEUR, CLAY. Reexaminar el significado del tiempo del profesorado. *Cooperación Educativa*; n. 69, p.26-31, 2006. Disponível em: <file:///C:/Users/Seven7/Downloads/archivoPDF.pdf>. Acesso : 06/05/2014.

LIMA, Luiz Cesar do Santos; SANTOS JUNIOR, Guataçara dos; MENDES, Piraju Borowski; MUNHOS, Jackson André. A satisfação do manutentor na área industrial: O caso em uma indústria frigorífica – In. **Revista Brasileira de Tecnologia Agroindustrial**. 2012 p. 757 – 769. Disponível em: <http://revistas.utfpr.edu.br/pg/index.php/rbta/article/view/901/860>. Acesso: 25/04/2014.

LOPES, Celi Aparecida Espasandin; **A probabilidade e a Estatística no Ensino Fundamental: Uma Análise Curricular**, (dissertação) Universidade Estadual de Campinas, SP. 1998. Disponível em: [file:///C:/Users/Seven7/Downloads/LopesC.A.E%20\(1\).pdf](file:///C:/Users/Seven7/Downloads/LopesC.A.E%20(1).pdf) .Acesso: 01/02/2015.

LOURENÇO FILHO, Manoel Bergström. **Educação Comparada**. 3ª ed, Brasília, INEP/MEC, 2004.

LUNA, Sergio Vasconcelos de. **Planejamento de Pesquisa**: uma introdução. São Paulo: EDUC, 2007.

LUNET, Nuno; SEVERO, Milton; BARROS, Henrique. **Desvio Padrão ou Erro Padrão**. Faculdade de Medicina da Universidade do Porto, Portugal, 2006. Disponível em: <http://www.scielo.oces.mctes.pt/pdf/am/v20n1-2/v20n1-2a08.pdf> Acesso: 03/02/2015.

MADDUX E. James. **SELF-EFFICACY: THE POWER OF BELIEVING YOU CAN** George Mason University - (in press). Snyder, C. R., & Lopez, S. J. (eds). *Handbook of positive psychology*. New York: Oxford University Press. September 30, 2000. Disponível em: <https://www.google.com.br/#q=self-efficacy+the+power+of+believing+you+can+james+e.+maddux> . Acesso: 10/05/2014.

MATTAR, FauzeNajib. **Pesquisa de marketing**. Edição Compacta. 3.ed. São Paulo: Atlas, 2001.

MECD. Ministerio de Educación, Cultura y Deporte. Espanha, 2014. **Sistema Estatal de Indicadores de la Educación (SEIE)**. Disponível em: <http://www.mecd.gob.es/inee/sistema-indicadores/Edicion-2014.html> Acesso: 15/08/2014.

MEDEIROS, Carlos Augusto de. **Estatística Aplicada a Educação**. Publicação PIC. Distrito Federal: Universidade de Brasília, 2007.

MEMORIA, José Maria Pompeu – Uma Breve História da Estatística – **EMBRAPA Informação e Tecnologia**, Brasília 2004. Disponível em:  
[http://www.im.ufrj.br/~lpbraga/prob1/historia\\_estadistica.pdf](http://www.im.ufrj.br/~lpbraga/prob1/historia_estadistica.pdf). Acesso 16/07/2014 e

MICHELOTTO, Regina Maria. **A Liberação do Acesso e a Extensão como Estratégias de Democratização da Universidade**: a experiência da Itália e do Brasil. 193 p. Tese (Doutorado em Educação) – Setor de Educação, Universidade Federal de São Carlos, São Carlos, 1999.

MINHOTO, Maria Angélica. Política de Avaliação da Educação Brasileira: limites e perspectivas. In: SOUZA, Ângelo Ricardo de; GOUVEIA, Andréa Barbosa; TAVARES, Tais Moura (orgs.). **Políticas Educacionais**: conceitos e debates. Curitiba: Appris, 2013, p. 163-187.

MIRSHAWKA, Victor. Estatística – **Campos de Aplicação e Possibilidade Profissional**. São Paulo, 1975

MOHALLEN, D. F; TAVARES, M; SILVA, P. L; GUMARAES, E. C; FREITAS, R. F. **Avaliação do Coeficiente de Variação como medida da precisão em experimentos com frangos de corte**. Scielo. Arquivo Brasileiro de Medicina Veterinária e Zootecnia, v. 60, n.2, p. 449 – 453, 2008. Disponível em:  
<http://www.scielo.br/pdf/abmvz/v60n2/a26v60n2.pdf>. Acesso: 02/03/2015.

MORAES, Carlos Mesquita. **Escala de Medidas, Estatística Descritiva e Inferência Estatística**. Escola Superior de Educação. Bragança, Portugal. 2005. Disponível em: <https://bibliotecadigital.ipb.pt/bitstream/10198/7325/1/estdescr.pdf>. Acesso: 02/02/2015.

NEVES, Maria Manuela Costa: Introdução a Inferência Estatística. **Instituto Superior de Agronomia (ISA)**. Lisboa, Portugal. 2009. Disponível em:  
<http://www.isa.utl.pt/dm/estat/estat/seb3.pdf>. Acesso: 02/02/2015.

NÓVOA, Antonio. (coord.). **Os professores e sua formação**. Lisboa: Dom Quixote, 1992.

OCDE. **Le rôle crucial des enseignants. Atirrer, former et retenir des enseignants de qualité**, 2005. Disponível:  
<http://www.oecd.org/dataoecd/39/43/34990974.pdf> Acesso: 29/03/2014.

\_\_\_\_\_. TALIS Tchenical Report. **OCDE**, Paris, 2008. Disponível em  
<http://www.oecd.org/dataoecd/16/14/44978960.pdf> Acesso em: 19/10/2012.

\_\_\_\_\_. OCDE Anual Report. **OCDE**, Paris, 2009. Disponível em  
<http://www.oecd.org/dataoecd/45/58/20946947.pdf> Acesso em: 19/10/2012.  
 Acesso em 20/10/2012.

\_\_\_\_\_. TALIS, 2014. Questionários dos diretores e professores (2012-13). Disponível em: <http://portal.inep.gov.br/web/talis/resultados>. Acesso: 10/01/2015.

OLIVEIRA, Andréia FróesGaluci. Testes estatísticos para comparação de médias. In. **Revista Eletrônica**. Artigo nº 76. Nutritine, v.5, nº 6, p. 777-788 Nov/Dez/2008. Disponível em: [http://nutritime.com.br/arquivos\\_internos/artigos/076V5N6P777\\_788\\_NOV2008\\_.pdf](http://nutritime.com.br/arquivos_internos/artigos/076V5N6P777_788_NOV2008_.pdf). Acesso: 07/02/2015.

OLIVEIRA, Kenny Delmonte; ALMEIDA, Keyla Lopes de; BARBOSA, Thiago Leite. **Amostragens probabilísticas e Não probabilísticas: técnicas e aplicações na determinação de amostras**. Programa de pós- graduação. UFES. ES. 2012. Disponível em: [files.wendelandrade.webnode.com.br/.../Amostragens%20probabilística](http://files.wendelandrade.webnode.com.br/.../Amostragens%20probabilística) Acesso: 10/02/2015.

OLIVEIRA, Tânia Modesto Veludo de. Escalas de Mensuração de Atitude: Thrstone, Osgood, Stapel, Likert, Guttman, Alpert. São Paulo. In. **Revista administração Online** [online], v.2, n.2, abril/maio/junho/2001. Disponível em: [http://www.fecap.br/adm\\_online/art22/tania.htm](http://www.fecap.br/adm_online/art22/tania.htm). Acesso: 16/05/2014.

PAIVA, Mirella Lopez Martini Fernandes; PRETTE, Zilda Aparecida Pereira Del. Crenças docentes e implicações para o processo de ensino-aprendizagem – (2009) In. **Revista Semestral da Associação Brasileira Escolar e Educacional (ABRAPEE)**. V. 13n. 01 jan/jun 2009 – 75-85. Disponível em: <http://www.scielo.br/pdf/pee/v13n1/v13n1a09.pdf>. Acesso: 29/04/2014.

PAMPLONA, Admur Severino; CARVALHO, Dione Lucchesi de: **A Lógica da Inferência Estatística e seu Ensino na Licenciatura em Matemática**. Ano: 20xx Disponível em: [http://www.sbembrasil.org.br/files/ix\\_enem/Comunicacao\\_Cientifica/Trabalhos/CC54855837687T.doc](http://www.sbembrasil.org.br/files/ix_enem/Comunicacao_Cientifica/Trabalhos/CC54855837687T.doc). Acesso: 01/02/2015.

PEREIRA, Gizele Adriana Maciel. Brasil e Argentina: **Um Estudo Comparado das Reformas Educacionais a partir do PISA 2000**. Dissertação (Mestrado) – Setor de Educação. UFPR, Curitiba, 2011, 247p.pp. 77-98.ISSN 0104-4060 In: [http://educa.fcc.org.br/scielo.php?script=sci\\_pdf&pid=S0104-40602003000300005&lng=pt&nrm=iso&tlng=pt](http://educa.fcc.org.br/scielo.php?script=sci_pdf&pid=S0104-40602003000300005&lng=pt&nrm=iso&tlng=pt). Acesso em 05/10/2013.

PONT, Beatriz; NUSCHE, Deborah; MOORMAN, Hunter.**Improving school leadership: policy and practice**. v. 1. OECD Publishing, 2008a. Disponível em: <http://www.oecd.org/education/preschoolandschool/44374889.pdf> Acesso: 03/04/2014.

REIS, Elizabeth. **Estatística Multivariada Aplicada**. Lisboa: Sílabo, 1997.

\_\_\_\_\_. **Estatística Descritiva. 7 ed.** Lisboa: Sílabo, 2008.

Reis, Elizabeth, MELO, Paulo, ANDRADE, Rosa. CALAPEZ, Tereza. Estatística Aplicada – Probabilidade, Variáveis Aleatórias, Distribuições Teóricas. Vol.1. 4ª ed. Lisboa: Edições Sílabo. 2003.

ROCHABRÚN Silva, Guillermo. **A Metodologia no Positivismo e no Marxismo: Teoria e Metodologia**. Série: Ensayos, nº 1, Área de Sociologia. Lima, Peru: dezembro de 1974.

SASS, Odair, MINHOTO, Maria Angélica Pedra. **Indicadores e Educação no Brasil: a avaliação como tecnologia**. InterMeio: Revista do Programa de Pós Graduação em Educação, Campo Grande, MS, v.17, n, 33, p 63-81, jan/jun. 2011 Disponível em: [http://www.constelaciones-rtc.net/02/02\\_11.pdf](http://www.constelaciones-rtc.net/02/02_11.pdf). Acesso: 03/09/2014

SAVIANI, Dermeval. *As Concepções Pedagógicas na História da Educação Brasileira*. Texto elaborado no âmbito do projeto de pesquisa “O espaço acadêmico da pedagogia no Brasil”, financiado pelo CNPq, para o “projeto 20 anos do Histedbr”. Campinas, 25 de agosto de 2005. Disponível em: [http://www.joinville.udesc.br/portal/professores/maria/materiais/3\\_Saviani\\_Concepcoes\\_Pedagogicas.pdf](http://www.joinville.udesc.br/portal/professores/maria/materiais/3_Saviani_Concepcoes_Pedagogicas.pdf) Acesso: 15/01/2014.

SCHWILLE, John; DEBELE, Martial (autores); SCHUBERT, Jane (colaboradora). *Former les enseignants: politiques et pratiques*. Paris: UNESCO/IPE, 2007. Disponível: <http://unesdoc.unesco.org/images/0015/001502/150261f.pdf>. Acesso em: 30/04/2014.

SENRA, Nelson. Uma Breve História das Estatísticas Brasileiras (1822 – 2002) – **Instituto Brasileiro de Geografia e Estatística (IBGE)** – Rio de Janeiro 2009. Disponível em: <http://biblioteca.ibge.gov.br/visualizacao/livros/liv42899.pdf>. Acesso: 15/07/2014.

SGANZERLA, Nelva Maria Zibetti. Aspectos relevantes da estatística e a evasão de estudantes no curso de graduação em Estatística da UFPR, 2001. 285p. Tese de Doutorado. Universidade Estadual Paulista Júlio de Mesquita. SP.

SILVA, Izabelle Fiorelle. **O Sistema Nacional de Avaliação: características, dispositivos legais e resultados**. Est. Aval. Educ. São Paulo, v. 21, n. 47, p. 427-448 set/dez. 2010. Disponível em: <http://www.fcc.org.br/pesquisa/publicacoes/eae/arquivos/1602/1602.pdf>. Acesso: 4/09/2014.

SILVA, Juliane Silveira Freire da. **Modelagem de Equações Estruturais: apresentação de uma metodologia** – Dissertação (mestrado) UFRG. Escola de Engenharia. Porto Alegre 2006. Disponível: <http://www.lume.ufrgs.br/bitstream/handle/10183/8628/000583213.pdf?sequence=1>. Acesso dia 03/05/2014.

SIEGEL, Sidney. **Estatística Não-Paramétrica**. São Paulo: Mc Graw-Hill, 1975.

SIPRAKI, Robson; SALES, Rossana Lopes; PEREIRA, Gisele Adriana Maciel. O uso da estatística na educação: A metodologia empregada na pesquisa TALIS.

**Congresso: 5º Encontro Internacional de Educação Comparada.** Belém PA. 2012. Disponível em: <http://www.sbec.org.br/evt2012/trab48.pdf>. Acesso: 10/02/2015.

SOUZA, Antônio Lisboa Leitão de. Estado e Administração Pública: tendências administrativas e de gestão. In: OLIVEIRA, Dalila Andrade; ROSAR, Maria de Fátima Félix (orgs). **Política e Gestão da Educação.** Belo Horizonte: Autêntica, 2002, p. 89-103.

SOUZA JUNIOR, Geraldo Lopes de. Elaboração e Análise de Indicadores. Manaus, AM: **Secretaria de Estado de Planejamento (SEPLAN)**, 200-. Disponível em: [http://www.seplan.am.gov.br/arquivos/download/arqeditor/apostila\\_indicadores.pdf](http://www.seplan.am.gov.br/arquivos/download/arqeditor/apostila_indicadores.pdf) Acesso: 12/06/2014.

SOUZA, Lanara Guimaraes de. Avaliação de políticas educacionais: contexto e conceitos em busca da avaliação pública. In: LORDÊLO, J; DAZZANI, MV. (orgs). **Avaliação educacional: desatando e reatando nós** [online]. Salvador: EDUFBA, 2009. Disponível em: <http://books.scielo.org/id/wd/pdf/lordelo-9788523209315-02.pdf>

TORRECILLA, F. Javier Murillo; GARRIDO, Martinez Cyntia: **Análisis de Datos Cuantitativo com SPSS em Investigación Socioeducativa** – Cuadernos de Apoyo 35. UAM ediciones. Madrid - Espanha. 2012

TROJAN, Rose Meri. **A questão da formação docente: o que dizem os informes.** Salamanca/Espanha: Universidad de Salamanca - Facultad de Educación - Departamento de Teoría e Historia de la Educación, 2009, 37 p. Relatório Técnico.

TROJAN, Rose Meri; SIPRAKI, Robson. O que dizem os professores sobre suas condições de trabalho no Brasil, na Espanha e no México: um estudo comparado a partir da pesquisa TALIS (OECD-2009). **InterMeio: revista do Programa de Pós-Graduação em Educação.** Campo Grande, MS, v. 19, n.38, p. 26-51, jul-dez. 2013. Disponível em: <http://www.intermeio.ufms.br/ojs/index.php/intermeio/article/view/347/294> Acesso: 24/10/2013.

TROJAN, Rose Mari; SÁNCHEZ, Miriam Mabel. Educação Comparada: Considerações teórico-metodológico no contexto da globalização. **III Congresso Nacional y II Encuentro Intercional de Estudios Comparados en Educación.** Buenos Aires, 25 a 27 de junho de 2009. Anais. Disponível em: <http://www.saece.org.ar/autores3.php> Acesso: 24/04/2013. Acesso: 24/03/2014.

UFPE -Universidade Federal de Pernambuco: Departamento de Estatística – **Graduação/Perfil Profissional.** Recife. 2015. Disponível em: [https://www.ufpe.br/de/index.php?option=com\\_content&view=article&id=291&Itemid=235](https://www.ufpe.br/de/index.php?option=com_content&view=article&id=291&Itemid=235). Acesso em: 05/02/2015.

USP. Universidade de São Paulo. Instituto de Ciências Matemáticas e de Computação. Notícias. **Qual a probabilidade de um estatístico ficar desempregado?** São Carlos: USP, 03/09/2013. Disponível em: [http://www.icmc.usp.br/Portal/Noticias/leituraNoticias.php?tipoNoticia=Ensino&id\\_noticia=240](http://www.icmc.usp.br/Portal/Noticias/leituraNoticias.php?tipoNoticia=Ensino&id_noticia=240) Acesso: 12/08/2014.

VALENTIM Jean Louis. **Ecologia Numérica**: uma introdução à análise multivariada de dados ecológicos. Rio de Janeiro: Interciência, 2000.

VARGAS, Glaucia Garcia Bandeira de: **A metodologia da Resolução de Problemas e o Ensino da Estatística no novo ano do Ensino Fundamental**. Dissertação (mestrado). UNIFRA, Santa Maria RS, 2013. Disponível em: [http://sites.unifra.br/Portals/13/Disserta%C3%A7%C3%B5es/2013/Glaucia%20Garcia%20Bandeira%20de%20Vargas\\_%20Dissesta%C3%A7%C3%A3o%20de%20Mestrado.pdf](http://sites.unifra.br/Portals/13/Disserta%C3%A7%C3%B5es/2013/Glaucia%20Garcia%20Bandeira%20de%20Vargas_%20Dissesta%C3%A7%C3%A3o%20de%20Mestrado.pdf) Acesso: 02/02/2015.

VALI, Lori: **Testes de Hipóteses Não Paramétricos** – Departamento de matemática e estatística. Apostila. (RFRGS) Porto Alegre 2008. Disponível em: [http://www.mat.ufrgs.br/~viali/estatistica/mat2282/material/apostilas/Testes\\_Nao\\_Parametricos.pdf](http://www.mat.ufrgs.br/~viali/estatistica/mat2282/material/apostilas/Testes_Nao_Parametricos.pdf). Acesso: 03/02/2015.

VIANNA, Heraldo Marelím. **Avaliações Nacionais em Larga Escala: análises e propostas. Estudos em Avaliação Nacional**. São Paulo. n.27, p. 41-76, jan-jun/2003. Disponível em: <http://www.fcc.org.br/pesquisa/publicacoes/eae/arquivos/1057/1057.pdf>. Acesso: 15/08/2014.